THE ECONOMIC JOURNAL

MARCH 1991

The Economic Journal, **101** (*March* 1991) 157–179
Printed in Great Britain

A VARIANCE DECOMPOSITION FOR STOCK RETURNS*

John Y. Campbell

Every profession has its occupational hazards. For economists, one of these is the risk that in the midst of a social occasion, we will be asked to forecast or interpret the movements of the stock market. In this lecture I shall try to do just that. Since this is a professional meeting, I shall give enough detail to glaze the eyes of any casual questioner. But I hope to be able to show that there is genuine intellectual interest, as well as popular appeal, in this issue.

It is important from the outset to distinguish the two parts of the request – to forecast the market, and to interpret the market. These tasks are quite different. To forecast the market means to predict price changes in the near future. To interpret the market means to explain, with the benefit of hindsight, why prices have changed in the way they have. This is something which the financial press does almost every day. But the financial press does not impose on itself the discipline of consistency; one day's explanation need not cohere logically with the next day's story. The task for academics is to find an interpretation which can consistently explain stock market movements over a long period of time.

Forecasting and interpretation are of course related. To give one example, the strict 'random walk' theory of stock prices implies that stock returns are unforecastable. The only correct answer to a request to forecast the market is to explain, patiently, why this task is impossible. The random walk theory also implies that all unexpected movements in stock prices must be due to news about future dividends; only one interpretation of market movements is permitted.[1]

In general, however, the relation between forecasting and interpretation is weaker. In this lecture I shall emphasise the distinction between changes in rational expectations of future dividends and changes in rational expectations of future returns. I call the former 'news about future dividends' and the latter

* This paper was delivered to the Royal Economic Society at Nottingham on March 27, 1990 as the H. G. Johnson Lecture. An earlier version is available as NBER Working Paper No. 3246. I am grateful to Rob Stambaugh for assistance with the data, to John Ammer for research assistance, and to Chris Gilbert, Pete Kyle, Masao Ogaki, Robert Shiller, and participants in the 1989 NBER Summer Institute workshop on New Econometric Methods in Financial Markets for helpful comments and discussion. I acknowledge financial support from the National Science Foundation and the Sloan Foundation.

[1] This statement is true only if one rules out 'rational bubbles', as discussed below.

'news about future returns'. I shall interpret unexpected stock returns by breaking them into components which are attributable to these two types of news. The relative importance of the two components depends not only on the forecastability of stock returns, but also on the time-series properties of the forecastable component of returns. If predictable returns are highly persistent, then a small degree of predictability will have a drastic effect on the interpretation of returns.

To make these ideas precise, I need a framework that relates stock prices, stock returns, and dividends. The standard present value formula is tractable only if expected returns are constant, which may be one reason why the academic literature has focused for so long on this unlikely special case. Recently, Robert Shiller and I have developed a log–linear approximation to the standard formula, which is tractable even when expected returns vary through time. It is also surprisingly accurate. This approximation provides the basic framework for the analysis.

I shall also use a regression of the stock return, measured over a short period such as a month, onto variables known in advance. Numerous papers have shown that such regressions have a modest but statistically reliable degree of explanatory power.[2] I shall augment this regression with other regression equations describing the evolution through time of the forecasting variables. The resulting vector autoregressive (VAR) system, in combination with the log-linear asset pricing framework, can be used to calculate the impact that an innovation in the expected return will have on the stock price, holding expected future dividends constant. The impact is the 'news about future returns' component of the unexpected stock return. The 'news about future dividends' component is obtained as a residual.

The approach used here is different from two others which have been popular in the recent literature. What I will call the contemporaneous regression approach regresses stock returns on contemporaneous innovations to variables which might plausibly affect the stock market (Cutler *et al.*, 1989*b*; Roll, 1988). This breaks the return into a component which is a reaction to measured news variables, and a residual (sometimes called 'noise'). But the reaction to measured news could occur either because traders' expectations of future dividends change, or because their expectations of future returns change. The contemporaneous regression approach does not distinguish these possibilities.

The univariate time-series approach studies the autocorrelation function of stock returns (Conrad and Kaul, 1988; Cutler *et al.*, 1989*a*, *b*; Fama and French, 1988*a*; Lo and MacKinlay, 1988; Poterba and Summers, 1988). The objective is to decompose prices into a 'transitory' and a 'permanent' component. The movements of the former are associated with changing rational expectations of returns, but the movements of the latter are not. The approach postulates an unobserved components model for the stock price, calculates the implications of the model for the autocorrelations of returns, and then uses observed

[2] A partial list of references is Campbell (1987), Cutler *et al.* (1989*a*), Fama and French (1988*b*, 1989), and Keim and Stambaugh (1986).

autocorrelations to estimate the model parameters. It is argued that if the observed autocorrelations are all zero, so that *ex post* stock returns are white noise, then this is evidence that expected returns are constant.

A practical problem is that the univariate time series often delivers only weak evidence against the hypothesis that all autocorrelations are zero.[3] This is because one loses power by forecasting returns using only past returns, ignoring all the other possible information variables. The autocorrelations of *ex post* returns can be very small even when expected returns are variable and highly persistent. The reason is that innovations in expected returns cause movements in *ex post* returns in the opposite direction; the resulting negative serial correlation in *ex post* returns tends to offset the positive serial correlation arising from persistent expected returns. In fact, it is possible to construct an example in which expected returns are variable and persistent, but *ex post* returns are white noise.[4] A related difficulty is that a strong assumption on the covariance of the two components is needed to identify the parameters of the model from the autocorrelations of returns. For both these reasons I argue that a multivariate approach is preferable to the exclusive focus on univariate autocorrelations.

This paper is based on the methods of Campbell and Shiller (1988*a*, *b*). Those papers decompose the variance of annual stock returns (and log dividend-price ratios) into components due to forecasts of cash flows and returns. Forecasting equations are estimated for dividend growth rates and log dividend-price ratios, rather than returns; but the forecasting system implies forecasts of returns, because the log stock return can be well approximated by a linear combination of dividend growth rates and log dividend-price ratios. In this paper I forecast returns explicitly and do not include dividend growth rates in the analysis. This has the advantage that I can work with higher-frequency monthly data without having to deal with seasonals in dividend payments.

The work of Kandel and Stambaugh (1988) is also relevant for this lecture. Kandel and Stambaugh examine the long-run implications of a low-order vector autoregression in returns and a set of forecasting variables. They are able to show that the low-order VAR can account for several long-run characteristics of the data, including the high R^2 statistics obtained in long-horizon regressions by Fama and French (1988*a*). However, Kandel and Stambaugh do not study the impact of expected return movements on unexpected returns.

Finally, Campbell (1990*a*) is a brief summary of some of the major results given here.

[3] This statement holds for papers which look at lower frequency returns and larger stocks; see for example Fama and French (1988*a*), Poterba and Summers (1988), and recent critiques by Jegadeesh (1989), Kim *et al.* (1989), Richardson (1989), and Stock and Richardson (1989). Conrad and Kaul (1988) and Lo and MacKinlay (1988) find strong evidence against white noise returns, but they are looking at high frequency weekly data and smaller stocks.

[4] Poterba and Summers (1988) claim that 'If market and fundamental values diverge, but beyond some range the differences are eliminated by speculative forces, then stock prices will revert to their mean. Returns must be negatively serially correlated at some frequency if "erroneous" market moves are eventually corrected' (pp. 27–8). The example given in section II below shows that this statement is not generally true.

The organisation of the paper is as follows. The next section sets up the basic framework which will be used to calculate the relation between unexpected returns and movements in expected returns. Section II describes and compares the univariate time-series approach and the VAR approach for decomposing the variance of stock returns. Section III reports empirical results for monthly U.S. data in the period 1927–88, and Section IV concludes.

I. EXPECTED RETURNS AND UNEXPECTED RETURNS

The basic equation used in this paper relates the unexpected real stock return in period $t+1$ to changes in rational expectations of future dividend growth and future stock returns. Writing h_{t+1} for the log real return on a stock held from the end of period t to the end of period $t+1$, and d_{t+1} for the log real dividend paid during period $t+1$, the equation is

$$h_{t+1}-\mathrm{E}_t h_{t+1} = (\mathrm{E}_{t+1}-\mathrm{E}_t)\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j}-(\mathrm{E}_{t+1}-\mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j h_{t+1+j}. \qquad (1)$$

Here E_t denotes an expectation formed at the end of period t, and Δ denotes a 1-period backward difference. The parameter ρ is a number a little smaller than one.

Equation (1) is best thought of as a consistency condition for expectations. If the unexpected stock return is negative, then either expected future dividend growth must be lower, or expected future stock returns must be higher, or both. To see why, consider an asset with fixed dividends whose price falls. Its dividend yield is now higher; this will increase the asset return unless there is a further capital loss. Capital losses cannot continue forever, so at some point in the future the asset must have higher returns.

The discounting at rate ρ in equation (1) means that an increase in stock returns expected in the distant future is associated with a smaller drop in today's stock price than is an increase in stock returns expected in the near future. To understand why this is, consider the arrival of news that stock returns will be higher ten periods from now. If the path of dividends is fixed, the stock price must drop to allow a rise ten periods from now. Most of the drop occurs today, but for nine periods there are smaller declines which are compensated by a higher dividend yield. These further declines reduce the size of the drop which is required today.

Formally, equation (1) follows from the log-linear 'dividend-ratio model' of Campbell and Shiller (1988*a*). This model is an appropriate framework because it allows both expected returns and expected future cash flows to affect asset prices. The model is derived by taking a first-order Taylor approximation of the equation relating the log stock return to log stock prices and dividends. The approximate equation is solved forward, imposing a terminal condition that the log dividend-price ratio does not follow an explosive process. Further details are given in the Appendix.

It will be convenient to simplify the notation in equation (1). Let us define $v_{h,t+1}$ to be the unexpected component of the stock return h_{t+1}:

$$v_{h,t+1} \equiv h_{t+1} - \mathrm{E}_t h_{t+1}.$$

Let us define $\eta_{d,t+1}$ to be the term in equation (1) which represents news about cash flows:

$$\eta_{d,t+1} \equiv (\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}.$$

Let us define $\eta_{h,t+1}$ to be the term in equation (1) which represents news about future returns:

$$\eta_{h,t+1} \equiv (\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=1}^{\infty} \rho^j h_{t+1+j}.$$

Then equation (1) can be rewritten as

$$v_{h,t+1} = \eta_{d,t+1} - \eta_{h,t+1}. \tag{2}$$

I.1. *A univariate AR(1) for the expected return*

A useful special case is that in which the expected stock return follows a univariate first-order autoregressive, or AR(1), process.[5] Let us define u_{t+1} to be the innovation at time $t+1$ in the one-period-ahead expected return: $u_{t+1} \equiv (\mathrm{E}_{t+1} - \mathrm{E}_t) h_{t+2}$. If the expected return follows a univariate time series process, then $\eta_{h,t+1}$ is an exact function of u_{t+1}. The AR(1) case is

$$\mathrm{E}_{t+1} h_{t+2} = \phi \mathrm{E}_t h_{t+1} + u_{t+1}, \tag{3}$$

which implies

$$\eta_{h,t+1} = \frac{\rho u_{t+1}}{1 - \rho\phi}. \tag{4}$$

Since ρ is a number very close to one, this equation says that a 1% increase in the expected return today is associated with a capital loss of about 2% if the AR coefficient is 0·5, a loss of about 4% if the AR coefficient is 0·75, and a loss of about 10% if the AR coefficient is 0·9. Poterba and Summers (1988, Appendix) give a similar result.

Equation (4) can also be used to calculate the ratio of the variance of news about future returns to the overall variance of unexpected returns. If the AR(1) model holds, this ratio satisfies

$$\begin{aligned}\frac{\mathrm{Var}(\eta_{h,t+1})}{\mathrm{Var}(v_{h,t+1})} &= (1-\phi^2)\left(\frac{\rho}{1-\rho\phi}\right)^2\left(\frac{\mathrm{R}^2}{1-\mathrm{R}^2}\right) \\ &\approx \left(\frac{1+\phi}{1-\phi}\right)\left(\frac{\mathrm{R}^2}{1-\mathrm{R}^2}\right),\end{aligned} \tag{5}$$

[5] It is important to note that this process does not restrict the size of the market's information set. In particular, there is no presumption that the relevant information set contains only the history of past asset returns. It is quite possible that a very large number of variables is useful in forecasting the asset return over the next period; the univariate AR(1) assumption merely restricts the way in which the next period's forecast is related to past forecasts. See Litterman and Weiss (1985) for discussion of the AR(1) model applied to the expected real return on short debt.

where R^2 is the fraction of the variance of stock returns which is predictable. If R^2 is 0·025, then the share of news about future expected returns in the variance of unexpected returns is 0·08 for $\phi = 0{\cdot}5$, 0·18 for $\phi = 0{\cdot}75$, and a startling 0·49 for $\phi = 0{\cdot}9$. These parameters are not unreasonable ones for monthly stock returns. An R^2 of 0·025 is quite modest, and a process with $\phi = 0{\cdot}9$ has a half-life of only a little more than six months. This example shows that movements in expected returns can be very important in explaining stock price volatility, even if the predictable component of the monthly stock return is small.

I.2. *Real returns and excess returns*

The discussion so far has been couched in terms of real stock returns. For many purposes it is more natural to work with excess stock returns over some short-term interest rate. If the log real interest rate is r_{t+1}, then the excess return is just

$$e_{t+1} \equiv h_{t+1} - r_{t+1}. \tag{6}$$

Since e_{t+1} is just the difference between two continuously compounded real returns, the price deflator cancels and it can equally well be measured as the difference between two log nominal returns.

It is straightforward to combine (6) and (1) to obtain

$$e_{t+1} - \mathrm{E}_t\, e_{t+1} = (\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=0}^{\infty} \rho^j r_{t+1+j},$$
$$- (\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=0}^{\infty} \rho^j e_{t+1+j}, \tag{7}$$

or in more compact notation,

$$v_{e,\,t+1} = \eta_{d,\,t+1} - \eta_{r,\,t+1} - \eta_{e,\,t+1}. \tag{8}$$

If the first two terms on the right hand side of (8) are treated as a composite residual, then (8) has exactly the same form as (2). Alternatively, one can work with the three-way decomposition of excess returns given on the right hand side of (8).

II. ALTERNATIVE APPROACHES TO VARIANCE DECOMPOSITION

One way to decompose the variance of stock returns is to examine the serial correlation of returns. I have called this the univariate time series approach. For example, if the expected real return follows the AR(1) process given in equation (3), then the *ex post* return follows an ARMA(1, 1) process whose ith autocovariance is:

$$\mathrm{Cov}\,(h_{t+1}, h_{t+1-i}) = \phi^{i-1}\left[\mathrm{Cov}\,(\eta_d, u) - \left(\frac{\rho}{1-\rho\phi} - \frac{\phi}{1-\phi^2}\right)\mathrm{Var}\,(u)\right]. \tag{9}$$

The autocovariances of *ex post* returns are all of the same sign and die off at rate ϕ; this is a property of the ARMA(1, 1) which is emphasised by Poterba and Summers (1988) and Cutler *et al.* (1989*a*).

Equation (9) reveals two difficulties with the univariate time series approach. First, if the autocovariances are not all zero then they identify ϕ and the term in square brackets in (9); but this is not generally enough to identify the innovation variance of expected returns, $\mathrm{Var}(u)$. For that one must make some assumption about the covariance $\mathrm{Cov}(\eta_d, u)$ between news about future dividends and shocks to expected returns. The assumption generally made is that the covariance is zero. But this is an arbitrary assumption, and I present evidence below that it is false.

Secondly, it is possible that all the autocovariances of stock returns are zero, even when expected returns are variable and persistent. In this case the univariate time series approach breaks down completely. The condition for this is

$$\mathrm{Cov}(\eta_d, u) = \left(\frac{\rho}{1-\phi\rho} - \frac{\phi}{1-\phi^2}\right)\mathrm{Var}(u). \tag{10}$$

Equation (10) can be satisfied with zero covariance $\mathrm{Cov}(\eta_d, u)$ between shocks to cash flows and shocks to expected returns, if the expected return follows a highly persistent process such that $\phi = \rho$. Alternatively, it can be satisfied with a positive covariance between shocks to cash flows and shocks to expected returns, and a less persistent expected return process which has $\phi < \rho$.

Of course, in practice it is unlikely that (10) will hold exactly. In the empirical work in the next section, I estimate the covariance between cash flow and expected return shocks to be negative in the U.S. stock market in 1927–88, so that one should find a predominance of negative autocorrelations in stock returns. But these autocorrelations are all small relative to their standard errors, which makes it hard for the univariate time series approach to give very definite results.

II.1. *The VAR approach*

Instead of focusing exclusively on the autocovariances of stock returns, I model the stock return as one element of a vector autoregression. I begin by working with real stock returns, and then modify the model to handle excess stock returns.

First, I define a vector $\mathbf{z}_{t+1}$ which has k elements, the first of which is the real stock return h_{t+1}. The other elements are other variables which are known to the market by the end of period $t+1$. Then I assume that the vector $\mathbf{z}_{t+1}$ follows a first-order VAR:

$$\mathbf{z}_{t+1} = \mathbf{A}\mathbf{z}_t + \mathbf{w}_{t+1}. \tag{11}$$

The assumption that the VAR is first-order is not restrictive, since a higher-order VAR can always be stacked into first-order (companion) form in the manner discussed by Campbell and Shiller (1988*a*). The matrix $\mathbf{A}$ is known as the companion matrix of the VAR.

Next I define a k-element vector $\mathbf{e}1$, whose first element is 1 and whose other elements are all 0. This vector picks out the real stock return h_{t+1} from the vector $\mathbf{z}_{t+1}$: $h_{t+1} = \mathbf{e}1'\mathbf{z}_{t+1}$, and $v_{h,t+1} = \mathbf{e}1'\mathbf{w}_{t+1}$. The first-order VAR generates simple multi-period forecasts of future returns:

$$\mathrm{E}_t h_{t+1+j} = \mathbf{e}1'\mathbf{A}^{j+1}\mathbf{z}_t. \tag{12}$$

It follows that the discounted sum of revisions in forecast returns can be written as

$$\eta_{h,t+1} \equiv (\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=1}^{\infty} \rho^j h_{t+1+j} = \mathbf{e1}' \sum_{j=1}^{\infty} \rho^j \mathbf{A}^j \mathbf{w}_{t+1}$$
$$= \mathbf{e1}' \rho \mathbf{A} (\mathbf{I} - \rho \mathbf{A})^{-1} \mathbf{w}_{t+1}$$
$$= \boldsymbol{\lambda}' \mathbf{w}_{t+1}, \tag{13}$$

where $\boldsymbol{\lambda}'$ is defined to equal $\mathbf{e1}'\rho\mathbf{A}(\mathbf{I}-\rho\mathbf{A})^{-1}$, a nonlinear function of the VAR coefficients.

Since $v_{h,t+1}$ is the first element of $\mathbf{w}_{t+1}$, $\mathbf{e1}'\mathbf{w}_{t+1}$, equations (13) and (2) imply that

$$\eta_{d,t+1} = (\mathbf{e1}' + \boldsymbol{\lambda}')\, \mathbf{w}_{t+1}. \tag{14}$$

These expressions can be used to decompose the variance of the unexpected stock return $v_{h,t+1}$, into the variance of the news about cash flow, $\eta_{d,t+1}$, the variance of the news about expected returns, $\eta_{h,t+1}$, and a covariance term.

In the VAR context there is no single measure of the persistence of expected returns. But one natural way to summarise persistence is by the variability of the innovation in the expected present value of future returns, relative to the variability of the innovation in the one-period-ahead expected return. Thus I define the VAR persistence measure P_h as

$$P_h \equiv \frac{\sigma(\eta_{h,t+1})}{\sigma(u_{t+1})} = \frac{\sigma(\boldsymbol{\lambda}'\mathbf{w}_{t+1})}{\sigma(\mathbf{e1}'\mathbf{A}\mathbf{w}_{t+1})}, \tag{15}$$

where $\sigma(x)$ denotes the standard deviation of x. Another way to describe the statistic P_h is to say that a typical 1 % positive innovation in the expected return will cause a P_h % capital loss on the stock. In the univariate AR(1) case, P_h would just equal $\rho/(1-\phi\rho)$, or approximately $1/(1-\phi)$.

Given the definitions of $\mathbf{A}$ and $\boldsymbol{\lambda}$ and a set of VAR estimates, it is straightforward to estimate the VAR persistence measure P_h and the variance decomposition for unexpected stock returns. The calculation of standard errors is a little more difficult. The approach I take is to treat the VAR coefficients, and the elements of the variance–covariance matrix of VAR innovations, as parameters to be jointly estimated by Generalised Method of Moments (Hansen, 1982). The GMM parameter estimates are numerically identical to standard OLS estimates, but GMM delivers a heteroskedasticity-consistent variance–covariance matrix for the entire set of parameters (White, 1984). Call the entire set of parameters $\boldsymbol{\gamma}$, and the variance–covariance matrix $\mathbf{V}$.

Any statistic such as the share of the variance of unexpected returns which is attributed to news about future expected returns can be written as a nonlinear function $f(\boldsymbol{\gamma})$ of the parameter vector $\boldsymbol{\gamma}$. Then I estimate the standard error for the statistic in standard fashion as $\sqrt{[f_\gamma(\boldsymbol{\gamma})'\mathbf{V}f_\gamma(\boldsymbol{\gamma})]}$.[6]

[6] This procedure is very similar to that used in Campbell and Shiller (1988*a*, *b*). However Campbell and Shiller treated the variance–covariance matrix of the variables in the VAR as fixed, and computed standard errors allowing only for sampling error in the VAR coefficients. Their procedure can be thought of as giving a standard error for a sample variance decomposition, while the procedure used here gives a standard error for a population variance decomposition.

The VAR approach can be used to analyse excess stock returns as well as real returns. In this case the basic equation is (8) rather than (2). If the first two terms on the right hand side of (8) are treated together, then equations (11) to (14) hold with e_{t+1} replacing h_{t+1}. The components of the excess stock return should be interpreted as 'news about future excess returns'.

It is more interesting to treat all the terms on the right hand side of (8) separately. The vector $\mathbf{z}_{t+1}$ must then include the excess return e_{t+1} as its first element, and the real interest rate r_{t+1} as its second element. Define $\mathbf{e2}$ to be a k-element vector whose second element is 1, with all other elements zero. Then 'news about future excess returns' is $\eta_{e,t+1} = \boldsymbol{\lambda}'\mathbf{w}_{t+1}$, with $\boldsymbol{\lambda}'$ defined as before. 'News about future real interest rates' is $\eta_{r,t+1} = \boldsymbol{\mu}'\mathbf{w}_{t+1}$, where

$$\eta_{r,t+1} = \mathbf{e2}'(\mathbf{I}-\rho\mathbf{A})^{-1}\mathbf{w}_{t+1} = \boldsymbol{\mu}'\mathbf{w}_{t+1}, \tag{16}$$

where $\boldsymbol{\mu}'$ is defined as $\mathbf{e2}'(\mathbf{I}-\rho\mathbf{A})^{-1}$. The residual 'news about future dividends', $\eta_{d,t+1}$, is given by

$$\eta_{d,t+1} = (\mathbf{e1}'+\boldsymbol{\lambda}'+\boldsymbol{\mu}')\,\mathbf{w}_{t+1}. \tag{17}$$

Separate persistence measures can be defined for the expected real interest rate and the expected excess return. For the real interest rate, the persistence measure is

$$P_r \equiv \frac{\sigma(\boldsymbol{\mu}'\mathbf{w}_{t+1})}{\sigma(\mathbf{e2}'\mathbf{A}\mathbf{w}_{t+1})}, \tag{18}$$

while the persistence measure for the excess return, P_e, is given by the formula in equation (15).

III. APPLICATION TO THE U.S. STOCK MARKET

For the sake of comparability with previous work, I use a standard data set here. I study the behaviour of the value-weighted New York Stock Exchange Index, as reported by the Center for Research in Security Prices (CRSP) at the University of Chicago.[7] The data set runs from 1926 to 1988, but I reserve the first year for lags so that my full sample period is 1927 to 1988. I deflate the nominal return on the index using the Consumer Price Index reported in Ibbotson Associates (1989).

The forecasting variables I use for the stock return are the lagged stock return, the dividend-price ratio, and the 'relative bill rate', the difference between a short-term Treasury bill rate and its one-year backward moving average. The lagged stock return is included because the stock return forecasting equation will be one equation of a VAR system. The dividend-price ratio is included, following Campbell and Shiller (1988*a*, *b*) and Fama and French (1988*b*), because it should reflect any changes that may occur in future expected returns.[8] The ratio is measured as total dividends paid over the previous year, divided by the current stock price.

[7] I obtain similar results using the CRSP equal-weighted index. The forecasting variables used here predict returns on large stocks as well as small stocks, so the choice of index is not critical.

[8] To see this, consider the *ex ante* version of equation (A 2) in the Appendix.

Table 1

Basic VAR Results for Real Stock Returns

Dependent variable	h_t (SE)	$(D/P)_t$ (SE)	$rrel_t$ (SE)	R^2	Joint significance	Stability
A: 1927:1–1988:12						
h_{t+1}	0·107	0·331	−0·424	0·024	0·018	0·111
	(0·063)	(0·283)	(0·195)			
$(D/P)_{t+1}$	−0·007	0·963	0·018	0·937	0·000	0·093
	(0·005)	(0·028)	(0·010)			
$rrel_{t+1}$	0·007	−0·040	0·669	0·450	0·000	0·025
	(0·005)	(0·016)	(0·061)			
B: 1927:1–1951:12						
h_{t+1}	0·142	0·483	0·926	0·028	0·183	0·533
	(0·091)	(0·466)	(0·712)			
$(D/P)_{t+1}$	−0·012	0·935	−0·033	0·901	0·000	0·309
	(0·007)	(0·045)	(0·041)			
$rrel_{t+1}$	0·005	−0·019	0·309	0·092	0·122	0·271
	(0·006)	(0·026)	(0·161)			
C: 1952:1–1988:12						
h_{t+1}	0·048	0·490	−0·724	0·065	0·000	0·512
	(0·060)	(0·227)	(0·192)			
$(D/P)_{t+1}$	−0·001	0·980	0·034	0·959	0·000	0·627
	(0·003)	(0·011)	(0·009)			
$rrel_{t+1}$	0·013	−0·017	0·739	0·548	0·000	0·375
	(0·012)	(0·058)	(0·052)			

Notes: h is the log real stock return over a month, (D/P) is the ratio of total dividends paid over the previous year to the current stock price, and *rrel* is the one-month Treasury bill rate minus a one-year backward moving average. Standard errors and test statistics are corrected for heteroskedasticity. 'Joint significance' is the significance level for a test of the hypothesis that all regression coefficients are zero. 'Stability' is the significance level for a Chow test that all regression coefficients are the same when the full sample is split at the end of 1951, or when the subsamples are split at their midpoints.

The relative bill rate is included because many authors, including Fama and Schwert (1977) and Campbell (1987), have noted that the level of short-term interest rates helps to forecast stock returns. The short-term interest rate itself may be nonstationary over this sample period, so it needs to be stochastically detrended. The subtraction of a one-year moving average is a crude way to do this; the relative bill rate can also be written as a triangular moving average of changes in the short-term interest rate, so it is stationary in levels if the short rate is stationary in differences.[9] The short rate used is the one-month Treasury bill rate series from Ibbotson Associates (1989).

One problem which arises when interest rate data are used is that the behaviour of interest rates has changed over time. In particular, the Federal Reserve Board held interest rates almost constant for much of the period up to 1951, when a Federal Reserve Board–Treasury Accord allowed rates to move more freely. Accordingly, I split the 1926–88 sample at the end of 1951. This also allows a separate look at the data from the period around the Great

[9] Another recently popular way to detrend the interest rate is to use the yield spread between interest rates of two different maturities. The relative bill rate has at least as much forecasting power for stock returns as the long-short yield spread, which is insignificant when it is added to the equations reported below.

Table 2

Variance Decomposition for Real Stock Returns

VAR specification and time period	R_h^2 (Sig.)	Var (η_d) (SE)	Var (η_h) (SE)	$-2\text{Cov}(\eta_d, \eta_h)$ (SE)	Corr(η_d, η_h) (SE)	P_h (SE)
h, D/P, rrel						
1 lag, monthly						
A: 1927:1–1988:12	0·024	0·369	0·285	0·346	−0·534	4·772
	(0·018)	(0·119)	(0·145)	(0·046)	(0·127)	(2·247)
B: 1927:1–1951:12	0·028	0·437	0·185	0·378	−0·664	3·258
	(0·183)	(0·226)	(0·182)	(0·053)	(0·118)	(2·414)
C: 1952:1–1988:12	0·065	0·127	0·772	0·101	−0·161	5·794
	(0·000)	(0·016)	(0·164)	(0·153)	(0·256)	(1·469)
h, D/P, rrel						
6 lags, monthly						
A: 1927:1–1988:12	0·087	0·538	0·265	0·197	−0·261	3·972
	(0·004)	(0·181)	(0·162)	(0·121)	(0·203)	(2·253)
B: 1927:1–1951:12	0·129	0·661	0·118	0·222	−0·398	1·909
	(0·083)	(0·363)	(0·142)	(0·288)	(0·565)	(1·515)
C: 1952:1–1988:12	0·118	0·127	0·797	0·075	−0·118	4·100
	(0·000)	(0·035)	(0·175)	(0·165)	(0·269)	(1·112)
h, D/P, rrel						
4 lags, quarterly						
A: 1927:1–1988:4	0·162	0·334	0·497	0·170	−0·208	2·726
	(0·045)	(0·096)	(0·193)	(0·186)	(0·269)	(1·435)
B: 1927:1–1951:4	0·307	0·428	0·476	0·096	−0·106	1·856
	(0·024)	(0·195)	(0·166)	(0·236)	(0·290)	(0·820)
C: 1952:1–1988:4	0·213	0·158	0·916	−0·074	0·097	7·289
	(0·000)	(0·067)	(0·184)	(0·211)	(0·257)	(5·437)

Notes: R_h^2 is the fraction of the variance of monthly real stock returns which is forecast by the VAR system, and Sig. is the joint significance of the VAR forecasting variables. η_d and η_h represent news about future dividends and news about future returns respectively. They are calculated from the VAR system using equations (14) and (13). The three terms Var (η_d), Var (η_h), and $-2\text{Cov}(\eta_d, \eta_h)$ are given as ratios to the variance of the unexpected stock return v_h, so from equation (2) they add up to one. The persistence measure P_h is defined in equation (15). A typical 1 % positive innovation in the expected real return is associated with a P_h % capital loss on the stock.

Depression, which may behave quite differently from the postwar data (Kim *et al.* 1989).

III.1. *Basic results for real returns*

Panel A of Table 1 reports the basic first-order VAR which I will use to analyse the persistence of expected returns. The first three columns give the regression coefficients for the stock return forecasting equation, the dividend-price ratio forecasting equation, and the relative bill rate forecasting equation. Together, these coefficients form the VAR companion matrix **A**. Heteroskedasticity-corrected standard errors are reported in parentheses. The remaining columns of the table report the regression R^2 statistics, the joint significance levels of the VAR forecasting variables, and the significance levels for Chow tests of parameter stability.

The R^2 statistic for the stock return equation is only 2·4 % over the full sample; the forecasting variables are jointly significant at the 1·8 % level, but the lagged stock return and the dividend-price ratio are individually insignificant. The other two equations have R^2 of 94% and 45% respectively, since the dividend-price ratio and relative bill rate follow quite persistent processes. There is little evidence of cross-effects between these two variables; each one is quite well modelled as a univariate autoregression. There is some evidence of instability in the VAR system, particularly in the equation describing the relative bill rate.

The top row of Table 2 calculates the implications of the VAR estimates for the variance of unexpected returns. Slightly more than a third of the variance of unexpected returns is attributed to the variance of news about future cash flows, slightly less than a third is attributed to the variance of news about future returns, and the remainder is due to the covariance term. The correlation between shocks to expected returns and shocks to cash flows is about -0.5, indicating that good news about fundamental value tends to be associated with declines in expected future returns. This means that stock prices move more in response to cash flow news than they would if expected returns were constant.[10] Finally, the persistence measure P_h is a little less than 5, indicating that a 1% positive innovation in the expected return is associated with a 5% capital loss. This is the same effect as would be given by a univariate AR(1) process for the expected return with a coefficient of just under 0.8.

Panels B and C of Table 1, and the second and third rows of Table 2, report the same statistics for the subperiods 1927–51 and 1952–88. There is no evidence of parameter instability between the first and second halves of these subperiods.

In 1927–51 the forecasting variables are jointly significant for stock returns at only the 18% level. The problem is partly that stock returns have a high variance which moves with the level of the dividend-price ratio; this means that the White heteroskedasticity correction greatly increases the standard error on this variable. The interest rate variable is not a good forecaster before 1952, which is not surprising given the interest rate regime of this period. Nonetheless, even in 1927–51 the variance decomposition of unexpected returns attributes less than half the variance to the variance of news about cash flow. The remainder is attributed to the variance of news about expected returns or the covariance term.

In 1952–88 stock returns are more strongly forecastable; the forecasting equation has an R^2 of 6.5%, and both the dividend-price ratio and the relative bill rate are highly significant. Three-quarters of the variance of unexpected stock returns is now attributed to news about future expected returns, and only one eighth to news about cash flow. Once again the correlation between shocks to fundamentals and shocks to expected returns is estimated to be negative, but less so than before.

[10] Campbell and Kyle (1988) emphasise a similar finding for annual U.S. stock market data over the period 1871–1986. But for an alternative view, see Barsky and DeLong (1989).

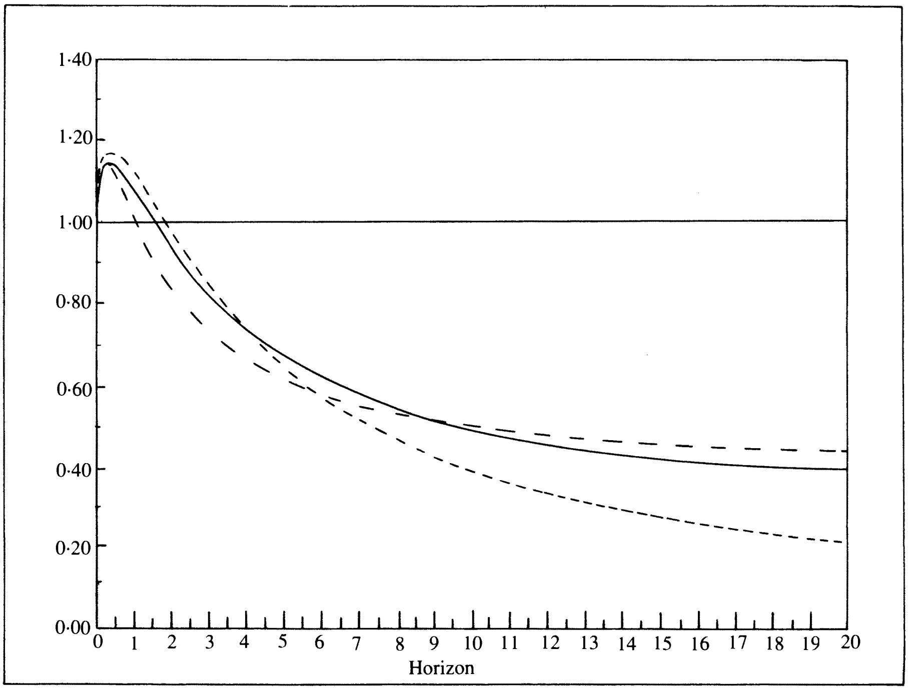


Fig. 1. Implied variance ratios. Solid line, 1927–88; dashed line, 1927–51; short dashed line, 1952–88. Variance ratios are calculated from the VAR(1) models estimated in Table 1.

III.2. *Univariate implications*

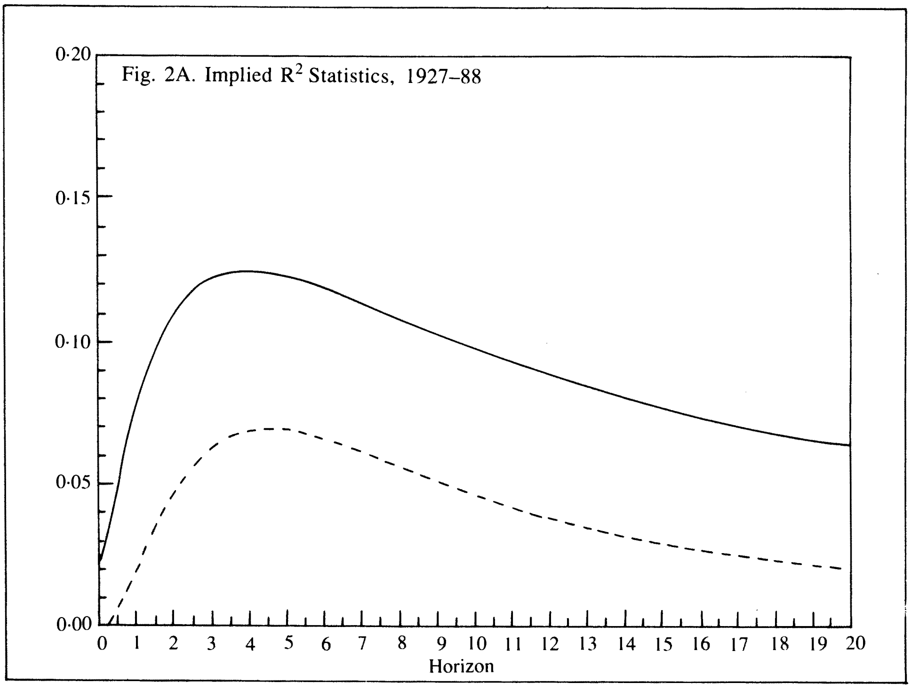
It is of some interest to calculate the univariate time-series implications of the VAR models estimated in Table 1.[11] Figure 1 shows the implied variance ratio statistic of Cochrane (1988) (see also Lo and MacKinlay, 1988 and Poterba and Summers, 1988), for horizons out to 20 years. These statistics are computed from the VAR estimates in Table 1, and not directly from the data.

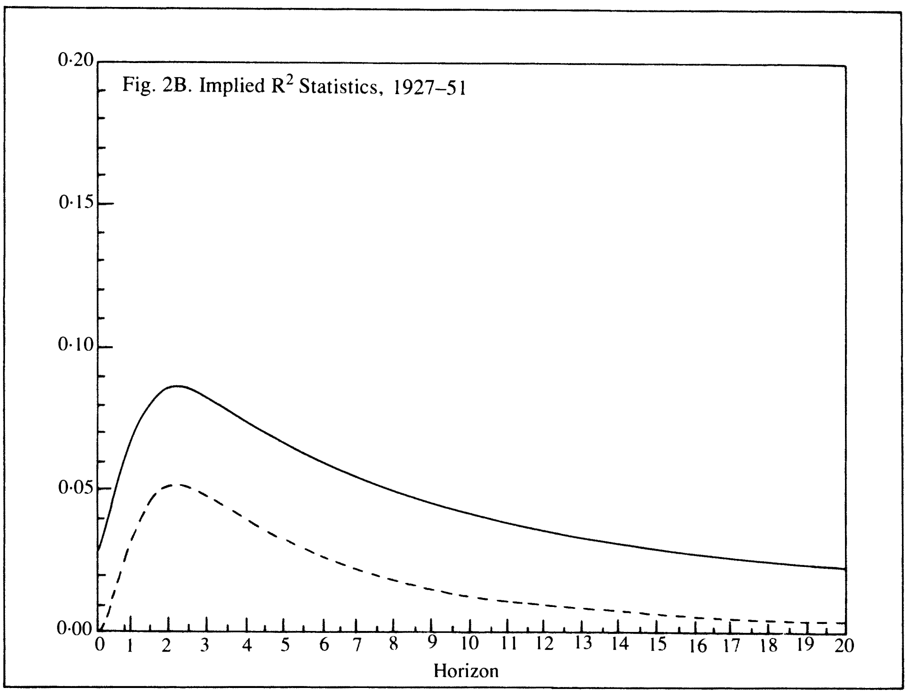
The variance ratio statistic $V(K)$ is defined as the ratio of the variance of K-period returns to the variance of 1-period returns, divided by K. This ratio will be one for white noise returns; it will exceed one for returns which are predominantly positively autocorrelated, and it will be below one when negative autocorrelations dominate. The ratio can be calculated directly from the autocorrelations of 1-period returns by using the fact that

$$V(K) = 1 + 2 \sum_{j=1}^{K-1} \left(1 - \frac{j}{K}\right) \mathrm{Corr}\,(h_t, h_{t-j}). \tag{19}$$

Implied variance ratios are shown for the periods 1927–88, 1927–51, and 1952–88. The general pattern is similar in the three periods. The variance ratios rise to a peak between 6 months and 1 year, and then decline steadily. The corresponding implied autocorrelations are positive for the first few

[11] Kandel and Stambaugh (1988) report a number of calculations of this type.





Figures 2A and 2B. For legend see opposite.

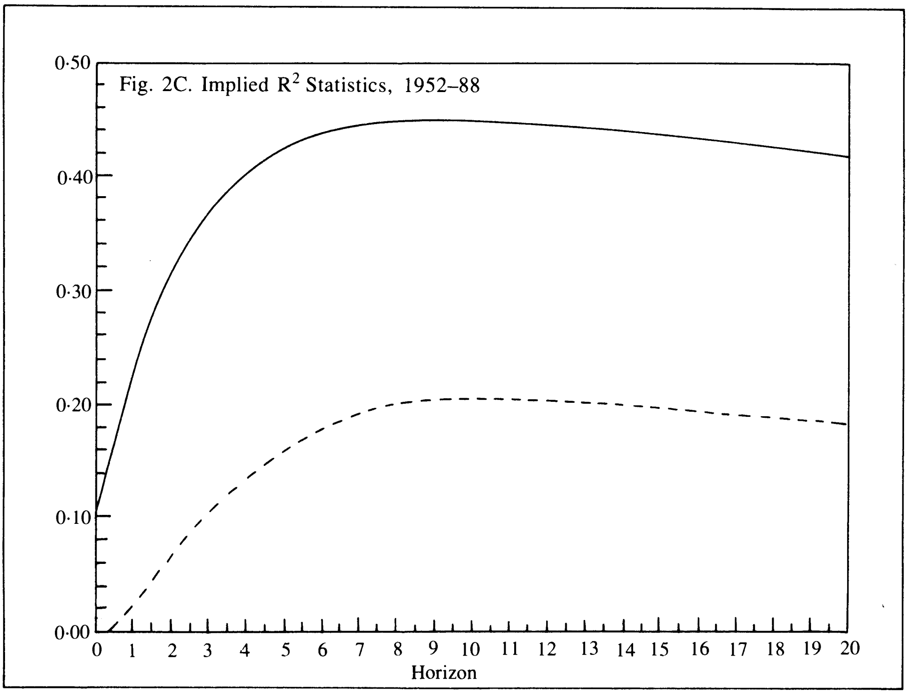


Fig. 2. Solid line, R^2 from regression of the K-year return on the VAR forecasting variables used in Table 1; dashed line, R^2 from regression of the K-year return on the lagged K-year return. K is measured on the horizontal axis. R^2 statistics are calculated from the VAR(1) models estimated in the respective panels of Table 1 for the matching time period.

months, and then negative; this is what Cutler (1989*a*) calls the 'characteristic autocorrelation function' of stock returns.

The postwar variance ratios differ from the prewar ones in that they rise further initially, and then decline more slowly. Beyond a horizon of 6 years, however, the postwar variance ratios lie below the prewar ones. They approach a limit of about 0·2, while the prewar limit is about 0·4. This long-run behaviour would of course be almost impossible to detect directly in the data.[12]

Fama and French (1988*a*) characterise the univariate behaviour of stock returns in another way. They regress the K-period stock return on the lagged K-period return. The Fama–French regression coefficient $\beta(K)$ is related to the variance ratio statistic by

$$\beta(K) = \frac{V(2K)}{V(K)} - 1. \tag{20}$$

The R^2 of the K-period regression is simply the square of $\beta(K)$. Figs. 2A, B, and C show the implied R^2 of Fama–French regressions for horizons out to 20 years. For comparison, the figures also show the implied R^2 from regressions of K-period returns on the variables used in the VAR system.

In all cases the R^2 statistics rise steeply from their initial values at a 1-month

[12] Kim *et al.* (1989) report that reliable direct evidence of mean reversion is found only in prewar data.

horizon to a peak at a horizon of several years.[13] In the full sample, the Fama–French R^2 peaks at about 7% at a horizon of 4 or 5 years, while the VAR R^2 peaks at about 13% at a horizon of 3 or 4 years. The prewar peak R^2 is 5% for the Fama–French regression and 9% for the VAR regression. In the postwar period, the R^2 statistics reach much higher peaks at longer horizons: almost 20% at 9 years for the Fama–French regression, and over 40% at 6 years for the VAR regression. In general the VAR R^2 statistics are about twice as high as the Fama–French R^2 statistics, which is an indication of the benefits obtainable from a multivariate rather than a univariate approach to stock returns.

III.3. *How robust are the results?*

The variance decomposition for stock returns is quite robust to changes in VAR lag length and data frequency. I support this claim in the lower panels of Table 2, where I give results for a monthly VAR with 6 lags and a quarterly VAR with 4 lags. The 6-lag monthly VAR gives almost the same results as before; the 4-lag quarterly VAR tends to give greater importance to the variance of news about future returns, and less importance to the covariance term, but is otherwise quite similar.

The variance decomposition is somewhat more sensitive to changes in the forecasting variables which are used. The critical variable appears to be the dividend-price ratio. When this variable is included in the VAR, the decomposition is much the same even if the relative bill rate is dropped from the system or replaced by the long–short yield spread. But when the dividend-price ratio is excluded from the system, the variance of news about future cash flows is given a more important role while the variance of news about future returns becomes less important and the covariance term is imprecisely estimated.

Fama and French (1989) suggest that the yield spread between low- and high-quality corporate bonds is a good substitute for the dividend-price ratio in forecasting stock returns. This appears to be true in the prewar period, but in the postwar sample the quality yield spread is insignificant for forecasting stock returns when it is added to the relative bill rate. Replacement of the dividend-price ratio by the quality yield spread causes a 1% reduction in the R^2 of the stock return forecasting equation. With this replacement, the VAR system attributes 1·686 (with a standard error of 0·719) of the variance of unexpected returns to the variance of news about cash flow, 0·649 (0·757) to the variance of news about future returns, and $-1{\cdot}334$ (1·465) to the covariance term. These estimates are much less precise than those reported in Table 2.

The variance decomposition can also be sensitive to the imposition of restrictions on the VAR system. For example, if one imposes the restriction that the expected stock return follows a univariate AR(1) process, the restricted expected return is primarily driven by the relative bill rate, with little weight

[13] This does not necessarily mean that econometric benefits are obtainable by working with long-horizon returns or further lags of the forecasting variables. In fact, if the VAR(1) model is correctly specified, then no further lags of the VAR variables are relevant for forecasting returns once the most recent values have been included.

Table 3

Monte Carlo Simulations

Time period	Sig. ($R_h^2 = 0$)	Var (η_d)/Var (v_h)	Sig. (Var (η_d)/Var (v_h) = 1)
A: 1927:1–1988:12			
Result in data	0·018	0·369	0·000
Empirical sig.:			
– estimated DGP	0·022	0·009	0·018
– unit root DGP	0·084	0·616	0·456
B: 1927:1–1951:12			
Result in data	0·183	0·437	0·013
Empirical sig.:			
– estimated DGP	0·225	0·123	0·175
– unit root DGP	0·433	0·877	0·772
C: 1952:1–1988:12			
Result in data	0·000	0·127	0·000
Empirical sig.:			
– estimated DGP	0·000	0·000	0·000
– unit root DGP	0·000	0·000	0·000

Notes: This table reports the results of a Monte Carlo experiment. The data generating process was the one estimated in Table 1, setting coefficients in the return forecasting equation to zero ('estimated DGP') or setting these coefficients to zero, the coefficient of D/P on lagged D/P equal to one, and the coefficient of *rrel* on lagged D/P equal to zero ('unit root DGP'). Each column of the table reports one statistic from Table 2. The result found in the data is reported, along with the fraction of 1000 runs which rejected the null more strongly than did the data. In the first column, the statistic is the joint significance of the VAR forecasting variables for monthly stock returns. In the second column, the statistic is the estimated share of news about future dividends in the variance of unexpected stock returns. In the third column, the statistic is the significance level for a Wald test of the hypothesis that this share equals one.

on the dividend-price ratio. (The system must use one or the other variable rather than both, because the two variables follow different AR(1) processes so no combination of them is well described by an AR(1) process.) The variance decomposition that results is similar to that obtained when the dividend-price ratio is dropped from the basic model. The univariate AR(1) restriction is not rejected in the prewar data, but is rejected at the 1·8% level in the postwar period.

A rather different concern about the variance decomposition is that it may be affected by small-sample bias. To address this concern, I report the results of a Monte Carlo study in Table 3. I use two alternative data generating processes (DGP's), both based on the models estimated in Table 1. The 'estimated DGP' simply takes the VAR system from Table 1 and sets the coefficients in the stock return forecasting equation to zero. Artificial data are generated by drawing standard normal random variables and feeding them through the system. The resulting data match the moments of the real data quite closely, except that the artificial stock returns are unforecastable and are moved entirely by news about future dividends, so that Var $(\eta_{d,t+1}) =$ Var $(v_{h,t+1})$.

The 'unit root DGP' adjusts the VAR coefficients from Table 1 in such a way that the dividend-price ratio follows a process with a unit root. The first row of the matrix **A** is set to zero as before, but also the (2, 2) element is set to

1 and the (3, 2) element is set to zero. This means that the matrix **A** has a unit eigenvalue.

The unit root DGP matches the data less well. (Dickey–Fuller tests reject the hypothesis that the dividend-price ratio has a unit root at the 1 % level or better in the full sample and both subsamples.) The reason for using this DGP is that unit roots in regressors cause well-known problems with standard statistical methods. The unit root DGP enables me to check whether the results reported above are robust to these types of problems.

The Monte Carlo study examines the small-sample behaviour of three statistics from Table 2. The first is the joint significance of the VAR variables for forecasting monthly stock returns, reported in the first column of Table 2. The second statistic is the ratio of the variance of news about future dividends to the total variance of unexpected stock returns, $\mathrm{Var}(\eta_{d,t+1})/\mathrm{Var}(v_{h,t+1})$, reported in the second column of Table 2. The third statistic is the significance level in a test of the hypothesis that this ratio equals one (computed from the standard error reported in the second column of Table 2). Table 3 repeats the results from Table 2, and then gives their empirical significance levels: the fraction of 1,000 Monte Carlo runs which generated statistics further from the null than the ones found in the data. This exercise is done separately for the full sample period and each subsample.

The estimated DGP reveals some finite-sample bias in the results of Table 2. There is only a very small bias in the predictability of one-period stock returns, but the test that the dividend variance ratio equals one is more seriously biased. Over the full sample, this test had an asymptotic significance level of better than 0·01 %, but it rejected more strongly in the artificial data 1·8 % of the time. In the first subsample, the asymptotic significance level is 1·3 %, but the empirical significance level is only 17·5 %. This bias is important, but it is not severe enough to change the general flavour of the empirical results.

The bias is greatly magnified when the unit root DGP is used. Over the full sample the empirical significance level of the test that the dividend variance ratio equals one is a startling 46 %, when the asymptotic significance level is less than 0·01 %. It is clear that the presence of a unit root in the dividend-price ratio has a devastating effect on the performance of the variance decomposition procedures. Even the unit root DGP, however, is unable to account for the results found in the postwar sample period. Not a single Monte Carlo simulation delivered more predictable stock returns, or dividend variance ratios further from one, than were found in the actual data. In the 1952–88 period, the evidence for predictable stock returns is overwhelming.

III.4. *Results for excess returns*

Table 4 reports a variance decomposition for excess returns, using equations (8), (16) and (17). The variance of unexpected excess returns is decomposed into the variance of news about dividends, $\eta_{d,t+1}$, the variance of news about real interest rates, $\eta_{r,t+1}$, the variance of news about future excess returns, $\eta_{e,t+1}$, and covariances among these shocks. The VAR system includes the excess return e_{t+1}, the real interest rate r_{t+1}, the dividend-price ratio, and the relative

Table 4

Variance Decomposition for Excess Stock Returns

VAR specification and time period	R_e^2 (Sig.)	Var (η_d) (SE)	Var (η_r) (SE)	Var (η_e) (SE)	-2Cov (η_d, η_r) (SE)	-2Cov (η_d, η_e) (SE)	2Cov (η_r, η_e) (SE)	P_r (SE)	P_e (SE)
e, r, D/P, rrel 1 lag, monthly									
A: 1927:1–1988:12	0·023	0·385	0·034	0·242	−0·018	0·322	0·035	1·639	4·376
	(0·051)	(0·127)	(0·012)	(0·145)	(0·052)	(0·059)	(0·052)	(0·793)	(2·295)
B: 1927:1–1951:12	0·029	0·452	0·049	0·119	0·056	0·281	0·043	1·529	2·381
	(0·257)	(0·234)	(0·027)	(0·173)	(0·124)	(0·161)	(0·058)	(0·833)	(2·167)
C: 1952:1–1988:12	0·065	0·137	0·012	0·734	−0·022	0·099	0·041	0·707	5·520
	(0·000)	(0·020)	(0·003)	(0·153)	(0·009)	(0·158)	(0·061)	(0·213)	(1·497)
e, r, D/P, rrel 4 lags, monthly									
A: 1927:1–1988:12	0·074	0·532	0·088	0·233	0·017	0·144	−0·015	2·230	3·617
	(0·011)	(0·190)	(0·041)	(0·148)	(0·107)	(0·137)	(0·091)	(1·152)	(2·113)
B: 1927:1–1951:12	0·105	0·762	0·122	0·125	0·221	−0·117	−0·114	1·934	1·960
	(0·293)	(0·474)	(0·068)	(0·179)	(0·287)	(0·596)	(0·241)	(1·060)	(1·763)
C: 1952:1–1988:12	0·109	0·129	0·034	0·764	−0·050	0·040	0·084	0·903	4·297
	(0·000)	(0·035)	(0·015)	(0·154)	(0·033)	(0·173)	(0·098)	(0·331)	(1·213)

Notes: R_e^2 is the fraction of the variance of monthly excess stock returns which is forecast by the VAR system, and Sig. is the joint significance of the VAR forecasting variables. η_d, η_r, and η_e represent news about future dividends, future real interest rates and future excess returns respectively. They are calculated from the VAR system using equations (13), (16), and (17). The terms Var (η_d), Var (η_r), Var (η_e), -2Cov (η_d, η_r), -2Cov (η_d, η_e), and 2Cov (η_r, η_e) are given as ratios to the variance of the unexpected excess return v_e, so from equation (8) they add up to one. P_r and P_e are the persistence measures for real rates and excess returns defined in equations (18) and (15) respectively.

bill rate. The first panel reports results for a 1-lag VAR, while the second panel gives results for a 4-lag VAR.[14]

In both panels the variance of news about future real interest rates is very small. The covariances between news about real interest rates and news about other variables are also small. It seems that news about real interest rates cannot account for large movements in stock prices.[15] One reason for this is that movements in the expected real rate are much less persistent than movements in the expected real stock return or the expected excess stock return. The persistence measure P_e in Table 4 is quite similar to the measure P_h in Table 2, whereas the persistence measure P_r for the expected real interest rate is much smaller. In general, the variance decomposition given in Table 4 is similar to that in Table 2.

IV. CONCLUSION

In this lecture I have argued that expected stock returns change through time in a fairly persistent fashion. The persistence of the expected return process means that changes in expected returns can cause sizeable capital gains and losses. Using monthly data on the New York Stock Exchange value-weighted index over the period 1927–88, I estimate that a typical 1 % increase in the expected return is associated with a capital loss of 4 or 5 %.

The variability and persistence of expected stock returns account for a considerable degree of volatility in unexpected returns. I estimate that over the full sample period, the variance of news about future cash flows accounts for only a third to a half of the variance of unexpected stock returns. The remainder of the stock return variance is due to news about future expected returns. This suggests that an economic explanation of stock market volatility must also be an explanation of short-term predictability in returns.

It is important to note that news about future returns is not independent of news about cash flows. Increases in future expected cash flows tend to be associated with decreases in future expected returns, a correlation which amplifies the volatility of stock returns. This finding is very similar to the 'overreaction to fundamentals' reported by Campbell and Kyle (1988).

Similar results obtain in the subperiods 1927–51 and 1952–88. However, the importance of expected return variation is greater in the postwar period. This contrasts with the fact that univariate analyses typically show greater evidence of mean reversion in stock prices before World War II.

The results apply almost equally well to real stock returns, and to stock returns measured in excess of a short-term Treasury bill rate. The variability of news about future excess stock returns is much greater than the variability of news about future real interest rates, and the latter has only a relatively small impact on stock returns. The reason for this is that the variables which forecast real interest rates are not highly persistent. This casts doubt on explanations of

[14] Since there are now 4 variables in each equation, there are more parameters to be estimated for any given lag length. A 6-lag VAR has 24 parameters in each equation and 106 parameters in all, which caused computational problems.

[15] For a similar argument that movements in real interest rate differentials do not account for real exchange rate movements, see Campbell and Clarida (1987).

variation in expected real stock returns which rely primarily on movements in real interest rates.[16]

Several caveats are worth mentioning. First, the results summarised above are point estimates with fairly wide confidence intervals. Both asymptotic standard errors and the results of a small Monte Carlo study show that there is only weak evidence for stock return predictability in the prewar period. The evidence that returns are predictable is overwhelming only in the period after 1952.

Secondly, the results are dependent on a particular specification of the information set which agents use to forecast stock returns. The variance decomposition alters considerably if the dividend-price ratio is excluded from the set of forecasting variables. It does not seem to be much affected by other changes in the information set, in particular by increases in VAR lag length or changes in the interest rate variable.

More generally, if agents have more information than is captured by the VAR, this can affect the variance decomposition given here. Of course, extra information can only increase the 1-period predictability of returns; but it is conceivable that it could decrease the share of return variance which is attributed to news about future expected returns. This possibility can be ruled out in the special case of a univariate AR(1) process for the expected return, but this case does not seem to fit the data particularly well.

It is also important to note that the variance decomposition of this lecture cannot be given an unambiguous structural interpretation. Although I have referred to 'news about dividends' and 'news about future returns', the results here do not necessarily mean that there is one-way causality from dividends and expected returns to stock prices. In general these variables are all determined simultaneously.

The approach I have developed here can be extended in several ways. It is of course very important to try to find an economic explanation of changing expected returns. Economic models of expected returns must fit the persistence and correlation with dividends observed in U.S. stock market data. Their implications for stock price behaviour can be studied in detail using the log-linear asset pricing framework of this lecture.[17]

Another interesting extension is the application of the VAR systems used here to cross-sectional asset pricing. Intertemporal asset pricing models explain cross-sectional risk premia as arising from the covariances of individual asset returns with the return on invested wealth (the market portfolio) and with news about future returns on invested wealth. The VAR structure used here can be a way to give these models increased empirical content.[18]

Princeton University

[16] One recent example is Cecchetti *et al.* (1990).

[17] See Campbell and Hentschel (1990) for an investigation of the relation between changing volatility and changing expected returns within this framework.

[18] See Campbell (1990*b*) for a restatement of cross-sectoral asset pricing theory using the framework of this lecture.

APPENDIX: THE DIVIDEND-RATIO MODEL

The dividend-ratio model of Campbell and Shiller (1988*a*) is derived by taking a first-order Taylor approximation of the equation defining the log stock return, $h_{t+1} = \log(P_{t+1}+D_{t+1}) - \log(P_t)$. The resulting approximation is

$$h_{t+1} \approx k + \delta_t - \rho\delta_{t+1} + \Delta d_{t+1}, \tag{A 1}$$

where h_{t+1} and d_t have already been defined, δ_t is the log dividend-price ratio $d_t - p_t$, and p_t is the log real stock price at the end of period t. The parameter ρ is the average ratio of the stock price to the sum of the stock price and the dividend, and the constant k is a nonlinear function of ρ.[19] Equation (A 1) says that the return on stock is high if the dividend-price ratio is high when the stock is purchased, if dividend growth occurs during the holding period, and if the dividend-price ratio falls during the holding period.

Equation (A 1) can be thought of as a difference equation relating δ_t to δ_{t+1}, Δd_{t+1} and h_{t+1}. Solving forward, and imposing the terminal condition that $\lim_{i\to\infty} \rho^i \delta_{t+i} = 0$, Campbell and Shiller (1988*a*) obtain

$$\delta_t = \sum_{j=0}^{\infty} \rho^j (h_{t+1+j} - \Delta d_{t+j+j}) - \frac{k}{1-\rho}. \tag{A 2}$$

This equation says that the log dividend-price ratio δ_t can be written as a discounted value of all future returns h_{t+j} and dividend growth rates Δd_{t+j}, discounted at the constant rate ρ less a constant $k/(1-\rho)$. If the dividend-price ratio is high today, this will give high future returns unless dividend growth is low in the future. It is important to note that all the variables in (A 2) are measured *ex post*; (A 2) has been obtained only by the linear approximation of h_{t+1} and the imposition of a condition that δ_{t+i} does not explode as i increases.

However (A 2) also holds *ex ante*. If one takes expectations of equation (A 2), conditional on information available at the end of time period t, the left hand side is unchanged since δ_t is in the information set, and the right hand side becomes an expected discounted value. Using the *ex ante* version of (A 2) to substitute δ_t and δ_{t+1} out of (A 1), I obtain (1).[20]

In the empirical work in the paper, I use sample means to set $\rho = 0{\cdot}9962$ for monthly data, and $\rho = (0{\cdot}9962)^3$ for quarterly data. The results are not sensitive to variation in ρ within a plausible range.

REFERENCES

Barsky, Robert R. and DeLong, J. Bradford (1989). 'Bull and bear markets in the twentieth century.' NBER Working Paper 3171, November.

Campbell, John Y. (1987). 'Stock returns and the term structure.' *Journal of Financial Economics*, vol. 18, pp. 373–99, June.

[19] Equation (A 1) and the other formulas given here differ slightly from those in Campbell and Shiller (1988*a*, *b*) because the notation here uses a different timing convention. In this paper I define the time t stock price and conditional expectation of future variables to be measured at the end of period t rather than the beginning of period t. This conforms with the more standard practice in the finance literature.

[20] See Campbell and Shiller (1988*a*) for an evaluation of the quality of the linear approximation in equations (A 1) and (A 2).

—— (1990*a*). 'Measuring the persistence of expected returns.' *American Economic Review Papers and Proceedings*, vol. 80, pp. 43–7, May.
—— (1990*b*) 'Intertemporal asset pricing without consumption.' Unpublished paper, Princeton University.
—— and Clarida, Richard H. (1987). 'The dollar and real interest rates.' *Carnegie–Rochester Conference Series on Public Policy*, vol. 27, pp. 103–39, Fall.
—— and Hentschel, Ludger (1990). 'No news is good news: an asymmetric model of changing volatility in stock returns.' Unpublished paper, Princeton University.
—— and Kyle, Albert S. (1988). 'Smart money, noise trading, and stock price behavior.' NBER Technical Working Paper 71, October.
—— and Shiller, Robert J. (1988*a*). 'The dividend-price ratio and expectations of future dividends and discount factors.' *Review of Financial Studies*, vol. 1, pp. 195–228, Fall.
—— and —— (1988*b*). 'Stock prices, earnings, and expected dividends.' *Journal of Finance*, vol. 43, pp. 661–76, July.
Cecchetti, Stephen G., Lam, Pok-sang and Mark, Nelson C. (1990). 'Mean reversion in equilibrium asset prices.' *American Economic Review*, vol. 80, pp. 398–418, June.
Cochrane, John H. (1988). 'How big is the random walk in GNP?' *Journal of Political Economy*, vol. 96, pp. 893–920, October.
Conrad, Jennifer and Kaul, Gautam. (1988). 'Time-variation in expected returns.' *Journal of Business*, vol. 61, pp. 409–25.
Cutler, David M., Poterba, James M. and Summers, Lawrence H. (1989*a*). 'Speculative dynamics.' Unpublished paper, June.
——, —— and —— (1989*b*). 'What moves stock prices?', *Journal of Portfolio Management*, vol. 15, pp. 4–12, Spring.
Fama, Eugene F. and French, Kenneth R. (1988*a*). 'Permanent and temporary components of stock prices.' *Journal of Political Economy*, vol. 96, pp. 246–73, April.
—— and —— (1988*b*). 'Dividend yields and expected stock returns.' *Journal of Financial Economics*, vol. 22, pp. 3–25, October.
—— and —— (1989). 'Business conditions and expected returns on stocks and bonds.' *Journal of Financial Economics* vol. 25, pp. 23–49, November.
—— and Schwert, G. William. (1977). 'Asset returns and inflation.' *Journal of Financial Economics*, vol. 5, pp. 115–46.
Hansen, Lars Peter (1982). 'Large sample properties of generalized method of moments estimators.' *Econometrica*, vol. 50, pp. 1029–54, November.
—— and Singleton, Kenneth J. (1983). 'Stochastic consumption, risk aversion, and the temporal behavior of asset returns.' *Journal of Political Economy*, vol. 91, pp. 249–65, April.
Ibbotson Associates. (1989). *Stocks, Bonds, Bills, and Inflation: 1989 Yearbook*. Chicago.
Jegadeesh, Narasimhan (1989). 'On testing for slowly decaying components in stock prices.' Unpublished paper. Anderson Graduate School of Management, UCLA, March.
Kandel, Shmuel and Stambaugh, Robert F. (1988). 'Modelling expected stock returns for short and long horizons.' Working Paper 42–88, Rodney L. White Center for Financial Research, Wharton School, University of Pennsylvania.
Keim, Donald B. and Stambaugh, Robert F. (1986). 'Predicting returns in the stock and bond markets.' *Journal of Financial Economics*, vol. 17, pp. 357–90, December.
Kim, Myung Jig, Nelson, Charles R. and Startz, Richard (1989). 'Mean reversion in stock prices? A reappraisal of the empirical evidence.' Unpublished paper. University of Washington, May.
Litterman, Robert B. and Weiss, Laurence (1985). 'Money, real interest rates and output: a reinterpretation of postwar U.S. data.' *Econometrica*, vol. 53, pp. 129–56, January.
Lo, Andrew W. and MacKinlay, A. Craig (1988). 'Stock market prices do not follow random walks: evidence from a simple specification test.' *Review of Financial Studies*, vol. 1, pp. 41–66, Spring.
Poterba, James M. and Summers, Laurence H. (1986). 'The persistence of volatility and stock market fluctuations.' *American Economic Review*, vol. 76, pp. 1142–51, December.
—— and —— (1988). 'Mean reversion in stock prices: evidence and implications.' *Journal of Financial Economics*, vol. 22, pp. 27–59, October.
Richardson, Matthew (1989). 'Temporary components of stock prices: a skeptic's view.' Unpublished paper. Graduate School of Business, Stanford University.
Roll, Richard (1988). 'R^2'. *Journal of Finance*, vol. 43, pp. 541–66, July.
Stock, James H. and Richardson, Matthew (1989). 'Drawing inferences from statistics based on multi-year asset returns.' Unpublished paper. Kennedy School of Government, Harvard University, and Graduate School of Business, Stanford University, November.
White, Halbert (1984). *Asymptotic Theory for Econometricians*, Academic Press: Orlando, Florida.