Econometric Aspects of the Variance-Bounds Tests: A Survey

Christian Gilles
Board of Governors of the Federal Reserve System

Stephen F. LeRoy
University of Minnesota

We survey the variance-bounds tests of asset-price volatility, stressing the econometric aspects of these tests. The first variance-bounds tests of the present-value relation reported apparently striking evidence of excess volatility of asset prices. The statistical significance of the results, however, was, either marginal or, in the case of model-free tests, impossible to assess. Moreover, the tests were soon criticized for a number of biases. Various other tests of the present-value relation were later developed, avoiding in different degrees the econometric problems attending the first-generation tests. The majority of these second-generation tests also found excess volatility, though sometimes of borderline statistical significance. This finding of excess volatility is robust and is difficult to explain within the representative-consumer, frictionless-market model.

The variance-bounds debate-whether asset prices are more volatile than traditional models imply-has

We have perceived helpful comments from Marjorie Flavin, Wayne Joerding, Hashem Pesaran, Richard Porter, Gary Shea. Robert Shiller. Jon Sonstelie, Douglas Steigerwald, and Kenneth West. This article should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or its staff.

The Review of Financial Studies 1991 Volume 4, number 4, pp. 753–791

engaged a number of economists over the past decade. At first glance, the variance-bounds controversy seems rooted in ideological considerations: as in debates in other areas of economics, it raises questions about the rationality of markets. In all such debates, the issue is the success of current neoclassical theory in explaining economic phenomena. Some argue for fundamental departures from the neoclassical program.

Fortunately, the variance-bounds debate has largely avoided the meaningless verbal exchanges that are typical of ideological disagreements. Genuine scientific differences, which are resolvable by further theoretical and empirical work, characterize this debate. As a result, progress has been rapid and a convergence of opinion appears to be occurring. While it is an exaggeration to characterize the variance-bounds debate as a model of how scientific work should proceed, the case could be made.

Econometric, not ideological or theoretical, issues have dominated the variance-bounds debate. As a consequence, evaluating the empirical evidence of excess volatility is difficult. Further, the debate is based on simulations in settings where straightforward analytical demonstrations are available. In many cases, these simulations obscure rather than clarify the main lines of development. As things stand, the time required to gain a foothold in this literature Is excessive.

In this article, we Introduce the variance-bounds debate and stress econometric issues-unlike LeRoy (1984, 1989) and West (1988b) who focus on theoretical issues. While the exposition is intended to be self-contained, readers should consult the original papers.

The theory and first-generation tests due to Shiller (1979, 1981) and LeRoy and Porter (1981) are reviewed in Section 1. Discussion of theory is brief to avoid duplication of LeRoy (1984, 1989) and West (1988b), and because those readers who wish to. master the econometric aspects of the debate will already be familiar with the underlying theory. In Section 2, we present Flavin's (1983) and Kleidon's (1986b) econometric critiques of the Shiller-LeRoy-Porter findings of excess volatility. Here theoretical analysis of the econometric properties of two-observation samples drawn from the autoregressive dividend model replaces the simulation methods used by Flavin and Kleidon. By relying on analytical methods, readers should see the intuition underlying Flavin and Kleidon's claim that the original tests are biased toward rejection. In Section 3, we turn to the "second-generation" variance-bounds tests, which also found excess volatility of asset prices using alternative econometric methods. We believe that these second-generation tests reinforce the conclusion based on the first-generation tests; asset-price volatility exceeds that predicted by the simplest present-value model. Granted this provisional conclusion, the question becomes why asset-price volatility is high. Here the literature has not arrived at even a provisional conclusion. Several possibilities are discussed in Section 4. Conclusions are offered in Section 5.

1. First-Generation Tests

1.1 The variance-bounds theorems

The variance-bounds inequalities are implications of the present-value relation. In our usage, this means that an asset's price is equal to the expected present value of the asset's payments, discounted at a constant rate.[1] Two questions arise: (i) Why focus on testing the present-value relation rather than consumption-based asset pricing more generally? (ii) Why test only the volatility implications of the present-value relation, rather than all its implications for the covariances between asset yields and prices? With regard to (i), whether yields of financial assets are sufficient to explain their prices is an important substantive question. If so, the principle of parsimony can be invoked in future modeling to justify ignoring the aggregate effects of risk aversion.

With regard to (ii), the same question-Why focus on one particular implication of the theory to the exclusion of all others?-could be directed equally well to tests of market efficiency that focus on return autocorrelations. The answer then would be that if, contrary to the present-value relation, future price changes can be inferred from past prices, analysts can recommend trading rules that outperform buy and hold in expected return. With regard to the variance-bounds tests, the alternative to the present-value relation might be that stock prices are driven by protracted waves of optimism or pessimism. The impli-

[1] The constancy assumption has caused misinterpretation: in the finance literature, it is not customary to include constancy of the discount rate in the definition of the present-value relation [e.g., Menon (1991)]. This element of generality In the finance literature's definition of the present-value relation is an outgrowth of the practice in elementary finance instruction of portraying net present value as the single unambiguously correct way to think about all problems in finance (in preference to such alternatives as internal rate of return as a criterion for capital budgeting, for example). In practice, the important problem in applying the present-value relation is to determine the magnitude of the adjustment for risk that is appropriate in the setting under consideration.

The problem of determining asset prices generally and that of determining what discount rate to use in the present-value relation are related tautologically. In general equilibrium, it makes no difference whether the analyst determines the price of financial assets directly, or indirectly from the present-value relation after determining the equilibrium discount factor. We follow here the practice, common strong nonfinancial economists, of analyzing asset pricing without explicit reference to the present-value relation, and then formulating the present-value relation as a falsifiable special case by assuming constancy of the discount rate. Empirical testing then determines whether the present-value relation appears in fact to be false.

The difference between these two usages is entirely semantic-our purpose is not to urge one framework or the other, but only to alert readers to a possible source of misunderstanding.

cation of excessive volatility is tested directly by variance-bounds tests. Such informal discussion suggests that variance-bounds tests have different power than regression-based tests under different alternative hypotheses and therefore merit independent study.

The main variance-bounds theorem is now reviewed. Define p^*_t, the ex post rational price of stock at date t (the term is Shiller's), as the present discounted value of future dividends.[2]

$$p^*_t \triangleq \sum_{i=1}^{\infty} \beta^i d_{t+i}. \tag{1}$$

The present-value relation says that actual stock price is the conditional expectation of ex post rational price

$$p_t = E(p^*_t \mid I_t), \tag{2}$$

where I_t is investors' information.

Under assumptions to be specified, the variance of ex post rational price and that of actual price are related by

$$V(p^*_t) = V(p_t) + \frac{\beta^2 V(r_t)}{1-\beta^2}, \tag{3}$$

where r_t represents excess return:

(Note our terminology: we use "return" and "rate of return" where others sometimes use "payoll" and "return," respectively.) To derive (3), note that (1) and (2) imply

$$p_t = \beta E(d_{t+1} + p_{t+1} \mid I_t)$$

$$p_t = \beta(d_{t+1} + p_{t+1} - r_{t+1}), \tag{5}$$

using (4). Now update (5) by one period and use the result to substitute out p_{t+1}, in (5). Iterating and using (i), there results

$$p_t = p^*_t - \sum_{j=1}^{\infty} \beta^j r_{t+j}.$$

Assume now that $V(r_t)$ isconstant over time, as occurs, for example under a linear covariance-stationary dividends process when information revelation is regular (as in the example below). Shifting the

[2] We let the symbol $\triangleq$ denote the defining equality.

rightmost term to the left side and taking variances (noting that the r_{t+j}'s are mutually uncorrelated), there results

$$V(p_t) + 2\,\mathrm{Cov}\left(p_t, \sum_{j=1}^{\infty} \beta^j r_{t+j}\right) + \frac{\beta^2 V(r_t)}{1-\beta^2} = V(p_t^*). \qquad (6)$$

Because rational forecasting implies in addition that p_t and r_{t+j} have zero covariance, (6) implies (3). Dropping subscripts, the latter equation becomes

$$V(p^*) = V(p) + \frac{\beta^2 V(r)}{1-\beta^2}. \qquad (7)$$

Equation (7) underlies all the variance-bounds tests discussed in this paper.' It says that the variance of ex post rational price equals the sum of the variance of actual price and that of returns (where the latter is multiplied by a constant which. depends on the discount rate). If investors have little information about future dividends, then the variance of price is low. Returns, however, buffeted by the largely -unexpected realizations of dividends, are highly volatile. In the opposite case in which agents can accurately predict dividends into the distant future, the variance of price is almost as high as that of ex post rational price. Returns, on the other hand, are very smooth.

LeRoy and Porter outlined two types of volatility tests: bounds tests and orthogonality tests [the terms are due to Durlauf and Hall (1989)]. The simplest bounds test is

$$V(p) \leq V(p^*),$$

a direct consequence of (7). A less straightforward bounds test is the LeRoy-Porter lower bound on $V(p)$. Define H_t to be the information set consisting of current and past dividends alone and take $\hat{p}_t$ to be the price that would prevail under H_t:

$$\hat{p}_t \triangleq E(p_t^* \mid H_t).$$

Then, assuming that I_t is at least as informative as H_t, the rule of iterated expectations implies that

$$\hat{p}_t = E(p_t \mid H_t),$$

but the conditional expectation of any random variable x is less volatile than x itself, implying

$$V(\hat{p}_t) \leq V(p_t)$$

or, assuming stationarity,

$$V(\hat{p}) \leq V(p). \qquad (8)$$

This lower bound on $V(p)$ is of interest primarily because it is equivalent to West's inequality on return volatility, discussed below.

The simplest orthogonality test of asset-price volatility determines whether $V(p^*)$, $V(p)$, and $V(r)$ are consistent with (7). If instead $V(p)$ and $V(r)$ are too large to be consistent with (7), the orthogonality of price and returns is rejected in favor of the hypothesis that these variables are negatively correlated [in view of (6), if $V(p_t) + \beta^2 V(r_t)/(1 - \beta^2)$ exceeds $V(p_t^*)$, the covariance term must be negative].

1.2 An example

An extended example will facilitate understanding of the variance-bounds theorems and will aid in the exposition of the sampling problems that attend their empirical implementation. Suppose that dividends follow a first-order Markov process

$$d_{t+1} = \lambda d_t + \epsilon_{t+1}, \tag{9}$$

where $E(\epsilon_t) = 0$, $V(\epsilon_t) = 1$, $|\lambda| < 1$, and the ϵ_t are serially independent. Here the mean of the dividend process is set to zero since a nonzero mean would drop out of all variance expressions. Further, agents are assumed to have sufficient information to forecast dividends perfectly up to m periods in advance. However, agents have no information about dividends beyond $t + m$ other than the information conveyed by dividends up to $t + m$. This all-or-nothing specification is restrictive if taken literally; its advantage is to provide a simple parameterization of the amount of information available to investors. As a consequence, comparison of actual and ex post rational stock prices is very easy, since the ex post rational stock price process is a limiting case ($m = \infty$) of the class of actual stock price processes.

The present-value relation yields

where p_t^m is the actual price of stock. Using (9) to form expectations and simplifying, (10) becomes

$$p_t^m = \frac{\beta\lambda}{1 - \beta\lambda} d_t + \sum_{i=1}^{m} \frac{\beta^i \epsilon_{t+i}}{1 - \beta\lambda}. \tag{11}$$

We now can compute the variance of p_t^m:

$$V(p_t^m) = \frac{\beta^2\lambda^2}{(1 - \lambda^2)(1 - \beta\lambda)^2} + \frac{\beta^2 - \beta^{2(m+1)}}{(1 - \beta^2)(1 - \beta\lambda)^2}. \tag{12}$$

$V(p_t^m)$ rises with m toward an asymptote,

$$V(p_t^\infty) = \frac{\beta^2(1 + \beta\lambda)}{(1 - \lambda^2)(1 - \beta\lambda)(1 - \beta^2)}, \tag{13}$$

which is just the variance of the ex post rational stock price.

Now we evaluate the variance of the one-period returns. Equation (5) implies that

$$r_{t+1}^m = d_{t+1} + p_{t+1}^m - \beta^{-1}p_t^m.$$

Substituting (11) for p_t^m and p_{t+1}^m and taking the variance, we get

$$V(r_t^m) = \frac{\beta^{2m}}{(1 - \beta\lambda)^2}. \tag{14}$$

Notice that $V(p_t^\infty)$, $V(p_t^m)$, and $V(r_t^m)$ as given by (13), (12), and (14)) respectively, satisfy

$$V(p_t^\infty) = V(p_t^m) + \frac{\beta^2 V(r_t^m)}{1 - \beta^2},$$

agreeing with (7).

1.3 Implementation

Shiller's (1981) implementation of an operational test of the inequality $V(p_t) \leq V(p_t^*)$ was simple and direct-more so than that of LeRoy and Porter, as will be seen below. To correct for trend,, Shiller divided through by constant growth rate trends. To understand Shiller's resolution of the problem that p_t^* is not observable, notice that the ex post rational price (1) is the solution to the recursion

$$p_t^* = \beta(p_{t+1}^* + d_{t+1}) \tag{15}$$

that satisfies the condition

$$\lim_{t\to\infty} \beta^t p_t^* = 0. \tag{16}$$

Shiller simply replaced p_t^* by the solution to (15) that satisfies instead the, terminal condition

$$p_T^* = \frac{1}{T}\sum_{t=1}^{T} p_t. \tag{17}$$

Denote by $p_{t|\bar{p}}^*$ the observable version of p_t^* generated by (15) and (17).

Shiller constructed sample estimates of the variances of p_t and $p_{t|\bar{p}}^*$ in the usual way as the average squared deviation of each variable from its own mean. Using a century-long data set, Shiller found that the standard deviation of actual stock prices exceeded that of the ex post rational stock prices by a factor of 5.59. Although no significance tests were reported, he interpreted this result as constituting rejection of the variance-bounds inequalities.

As noted in Section 1.1, LeRoy and Porter defined their null hypothesis from the equality version (7) of the variance-bounds relation. They assumed that dividends and stock prices; adjusted for trend as described below, were generated by a covariance-stationary bivariate linear process, with parameters restricted by (7). A simplified and intuitive version of the LeRoy-Porter implementation goes as follows: (i) estimate a linear autoregressive model for dividends and estimate β as the reciprocal of 1 plus the average rate of return on stock; (ii) estimate $V(p_t^*)$ by applying the present-value relation (1) directly to the model for dividends (thus avoiding the problem that p_t^* is unobservable); (iii) estimate $V(r_t)$ from the observable series of one-period returns; and (iv) estimate $V(p_t)$ from a linear model for p_t. It was found that the point estimate of each of the two terms on the right-hand side of (7) by itself exceeds the term on the left-hand side, indicating rejection of the variance bounds. However, the associated confidence intervals based on asymptotic distributions were SO wide that the rejection of (7) was of borderline statistical significance.

LeRoy and Porter corrected for trend, not by implementing a . mechanical trend adjustment, but by reversing the effect of the variables which cause the trend: inflation and corporate retained earnings.[3] Assuming that the returns per unit of capital are covariance stationary, the LeRoy-Porter algorithm should produce stationary series for price and dividends. Instead, LeRoy and Porter noted (1981, p. 569) that the adjusted series showed some evidence of a downward trend, but the LeRoy-Porter implementation of the trend correction was flawed, as Kleidon (1986b, note 4) pointed out.[4]

[3] Specifcally,inflation was removed from the data by dividing a commodity price index into dividends, earnings, and prices. To reverse the effect of earnings mention, It was assumed that the value of stock equals the product of a quantity index representing the amount of physical capital held by firms and a price index representing the dollar value per unit of this capital. If It is assumed that the physical capital acquired with retained earnings is purchased at the same price as currently available capital is valued, and is equally productive, then data on retained earnings can be used to compute rccursively the price and quantity variables.

[4] In the 1988 vesion of this paper, we provided a corrected exposition of the algorithm and LeRoy and Parke Implemented It In the 1988 version of their 1990 paper, using the 112-year sample of Shiller instead of the LeRoy-Porter postwar sample. The downward trend in the supposedly trend-adjusted series that LeRoy and Parke reported was even more pronounced in the 112-year sample than that watch LeRoy and Porter had reported. We do not know the reason for this downward trend, but It dearly invalidates the LeRoy-Porter trend correction and renders It of secondary interest (except, perhaps. to researchers interested in investigating the cause of the downward trend, which is worthy of study in its own right). Because the trend correction failed at its assigned task, detailed description of it was deleted from this paper and that of LeRoy and Parke (the earlier versions of these papers are available from the authors).

2. The Critics

2.1 Flavin[5]

Flavin's (1983) paper made two criticisms of Shiller's econometric tests. The first was that both the variance of p_t, and that of p_t^* are estimated with downward bias in small samples. Further, the effect is more severe for p_t^* than for p_t, implying possible reversal of the empirical counterpart of the variance-bounds inequalities even if the present-value relation is true. The second is that Shiller's procedure for calculating an observable version of p_t^* also induces bias toward rejection.

To understand Flavin's first criticism, we review the elementary statistics involved in estimating the variance of a population. If a sample $\{x_1, x_2, \ldots, x_n\}$ is drawn from a common distribution with known mean μ, the average squared deviation around the population mean yields an unbiased estimator of the variance:

$$\frac{E\left[\sum_{i=1}^{n} (x_i - \mu)^2\right]}{n} = \sigma^2. \tag{18}$$

Note that it is not necessary to assume that the x_i are mutually uncorrelated.

If the mean must be estimated, however, the variance estimator constructed by substituting the sample mean for the population mean in (18) is biased toward zero. The standard remedy is to reduce the degrees of freedom by 1:

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}.$$

The resulting estimator is unbiased only under the additional assumption that the x_i are uncorrelated. Flavin pointed out that both p_t and p_t^* are autocorrelated positively. Therefore reducing degrees of freedom by 1 will not provide an unbiased estimator of the variance of either p_t or p_t^*. Further, p_t^* is more strongly autocorrelated than p_t, implying that the variance of p_t^* is estimated with a greater downward bias than that of p_t. Flavin's argument suggests that if the present-value model is rejected when the volatility statistic $V(p^*) - v(p)$ is negative, type I error (rejecting the null when it is true) will occur with high probability. Therefore, the rejection region should be chosen smaller. Ideally, one would prespecify the probability of type I error and choose the rejection region accordingly. Under model-free

[5] Kleidon's (1986a) paper, written contemporaneously with Flavin's (1983) paper, made Independently many of the same points as Flavin.

tests, however, this is impossible because the distribution of the volatility statistic is unspecified.

To illustrate these assertions, we consider the sampling distribution of the sample variance of p_t^* and p_t, based on samples of size 2, from the example in Section 1.2. The assumption of two-element samples is less restrictive than it might seem: a sample of size 2, but with observations separated by n, periods, is qualitatively similar to a consecutive sample of size n. This is so because for large n, p_{t+n} is virtually independent of p, and similarly for p_{t+n}^* and p_t^*. We follow Flavin in assuming that the discount rate β is known. One would like to avoid this restrictive assumption, but the calculations to be presented would be rendered intractable if sampling variability of the estimate of β were considered. It is also assumed that p_t^* is measured correctly, pending discussion below.

Assume, as in Section 1.2, that individuals can forecast dividends up to m periods in the future. The sample variance of p_t^m is given by

$$\hat{V}_n(p_t^m) = (p_{t+n}^m - \bar{p})^2 + (p_t^m - \bar{p})^2, \tag{19}$$

where $\bar{p} = (p_{t+n}^m + p_t^m)/2$. Expression (19), which' reflects the usual correction for degrees of freedom (the sum of squares is divided by the number of observations, 2, less 1) simplifies to

$$\hat{V}_n(p_t^m) = \frac{(p_{t+n}^m - p_t^m)^2}{2}.$$

We wish to take the expectation of the sample variance. under the assumption that dividends follow (9). To do so, use (11) to write p_{t+n}^m as

$$\begin{aligned} p_{t+n}^m &= \frac{\beta\lambda d_{t+n} + \Sigma_{i=1}^m \beta^i \epsilon_{t+n+i}}{1-\beta\lambda} \\ &= \frac{\beta\lambda(\epsilon_{t+n} + \cdots + \lambda^{n-1}\epsilon_{t+1} + \lambda^n d_t) + \Sigma_{i=1}^m \beta^i \epsilon_{t+n+i}}{1-\beta\lambda}. \end{aligned}$$

Using (11) again to obtain an expression for p_t^m, subtracting this from p_{t+n}^m, squaring, taking the expectation, and dividing by 2, we obtain

$$\begin{aligned} E\left[\hat{V}_n(p_t^m)\right] = &\left[\frac{\beta^2\lambda^2(\lambda^n-1)^2}{1-\lambda^2} + \sum_{i=1}^{n} (\beta\lambda^{n+1-i} - \beta^i)^2 \right. \\ &\left. + \sum_{i=n+1}^{m} (\beta^{i-n} - \beta^i)^2 + \sum_{i=m+1}^{m+n} \beta^{2(i-n)}\right] \\ &\times \left[2(1-\beta\lambda)^2\right]^{-1}, \end{aligned} \tag{20}$$

for $n \leq m$, and

$$E\left[\hat{V}_n(p_t^m)\right] = \left[\frac{\beta^2\lambda^2(\lambda^n - 1)^2}{1-\lambda^2} + \sum_{t=1}^{m} (\beta\lambda^{n+1-t} - \beta^t)^2 + \sum_{t=m+1}^{n} (\beta\lambda^{n+1-t})^2 + \sum_{t=n+1}^{m+n} \beta^{2(t-n)}\right] \times \left[2(1-\beta\lambda)^2\right]^{-1}, \tag{21}$$

for $n > m$, and

We now use these expressions for the expectation of the sample variance to verify some properties of the bias of the variance estimators. First, in large samples, the variance of p_t^m is estimated without bias:

$$\lim_{n\to\infty} E\left[\hat{V}_n(p_t^m)\right] = V(p_t^m),$$

for all m and λ [let n go to infinity in (21) and observe that the result coincides with (12)]. It makes sense for the sample variance of p_t^m (corrected for degrees of freedom) to be an asymptotically unbiased estimator of $V(p_t^m)$ since p_{t+n}^m is approximately independent of p_t^m when n is large, even when λ is high (but less than unity).

Second, we have as a special case $(m = \infty)$ of (20)

$$\lim_{n\to\infty} E\left[\hat{V}_n(p_t^*)\right] = V(p_t^*),$$

so the sample variance of p_t^* gives an asymptotically unbiased estimate of its population variance as well. With both actual and ex post rational stock price estimated without bias, there is no bias problem in large-sample tests.

For small samples and unrestricted m and λ, the sample variance of p_t^m underestimates its population variance for two reasons. First, with $\lambda > 0$, the positive autocorrelation of dividends induces a corresponding positive correlation between nearby values of p_t^m. The simple correction for degrees of freedom takes no account of this dependence. Second, since agents can foresee dividends up to m periods ahead, the dividend realizations upon which p_{t+n}^m is based overlap those on which p_t^m is based if $m > n$. Hence again bias results. If both of these conditions are absent (i.e., if l = 0 and m < n), the sample variance is unbiased for any sample size:

[6] The (unavoidably tedious) verification of these equations involves isolating the coefficient of d_t and each of the ϵ_{t+r}, squaring term-by-term, and using the facts (1) $E[(\epsilon_t)^2] = 1$; (ii) $E[(d_t)^2] = 1/(1-\lambda^2)$; and (iii) d_t, is independent of the ϵ_{t+r}.

[verify the equality of the appropriately restricted versions of (21) and (12)].

Even if these conditions are met, the test statistic of the variance-bounds inequality is biased toward rejection—$V(p_t^*)$ is estimated without bias, but with n finite, $V(p_t^*)$ is underestimated as a result of the second effect noted above.[7] Thus, rejection of the variance-bounds inequality is likely infinite samples even if the null hypothesis is true. We see that whether dividends are autocorrelated or not, and however many periods ahead agents can forecast dividends, the sample variances provide an unbiased estimate of the corresponding population variances only in large samples, and that in finite samples this problem is more severe for the ex post rational price than for the actual price.

We turn now to Flavin's second criticism of Shiller's empirical implementation of the variance-bounds tests: bias is induced in the test statistic when an observable proxy is substituted for the unobservable p_t^* series [see also Shea (1989)]. As noted in Section 1, Shiller generated an observable version $p_{t|\bar{p}}^*$ of the p_t^* series from the recursion (15), supplying the remaining degree of freedom by imposing terminal condition (17). For each t, $p_{t|\bar{p}}^*$ in general gives a biased estimate of p_t^*, conditional on the realization of the sample. To show this, compare $p_{t|\bar{p}}^*$ to the series $p_{t|T}^*$ constructed as the solution to the recursion (15) satisfying the terminal condition $p_{T|T}^* = p_T$. By (2), $p_{T|T}^* = E(p_T^* | I_T)$. And if at any date $t \leq T$, $p_{t|T}^* = E(p_t^* | I_T)$, then

$$p_{t-1|T}^* = \beta(d_t + p_{t|T}^*) = \beta[d_t + E(p_t^* | I_T)] = E(p_{t-1}^* | I_T),$$

where the last equality follows from (15). This proves that at each date t in the sample, $p_{t|T}^* = E(p_t^* | I_T)$, but if $p_{t|T}^*$ is an unbiased estimate of p_t^*, then $p_{t|\bar{p}}^*$ must be biased to the extent that the two series differ. The fact that $p_{t|T}^*$ *is* a better proxy for p_t^* than $p_{t|\bar{p}}^*$ is now well understood and more recent variance-bounds papers, such as Mankiw, Romer, and Shapiro (1985) and Durlauf and Hall (1989), use the unbiased series.

We just established that $p_{t|T}^*$ estimates p_t^* without bias, but this does not mean that the sample variance of $p_{t|T}^*$ estimates the expectation of the sample variance of p_t^* without bias.[8] To see that the sample variance of $p_{t|T}^*$ will not be closely related to the expected sample variance of p_t^*, it suffices to observe that (except in the trivial

[7] For example, (13) Implies that $V(p_t^*) = \beta^2/(1 - \beta^2)$, while (20) implies that $E[\hat{V}_n(p_t^*)] = \beta^2(1 - \beta^n)/(1 - \beta^2)$, which is smaller than $V(p_t^*)$.

[8] The latter, rather than the population variance of p_t^*, is the relevant parameter for comparison if we wish to serve the bias induced by incorrect measurement of p_t^*, under discussion here, from the general problem of small sample bias discussd above.

certainty case) the population variance of $p^*_{t|T}$ is not constant over time even if the population variance of p^*_t is constant. Because $p^*_{T|T} = p_T$, we have $V(p^*_{T|T}) = V(p_T) < V(p^*_T)$ by the variance-bounds theorem, but $\lim_{t\to\infty} V(p^*_{T-t|T}) = V(p^*_T)$ since the $p^*_{T-t|}$ series converges toward the true p^*_t going backward in time.[9]

It is clear intuitively why substituting the observable $p^*_{t|T}$ for the unobservable p^*_t will reduce the estimated variance of ex post rational price. The two variables are equal, except that in $p^*_{t|T}$ the innovations in dividends occurring after the end of the sample are set equal to zero, which can only reduce the variance of ex post rational price. Of course, the effect is strongest near the end of the sample. It follows that in long samples the downward bias induced by substituting $p^*_{t|T}$ for p^*_t will be negligible since at most dates $p^*_{t|T}$ will be very Close to p^*_t. For smaller samples, however, the effect may be far from negligible.

Flavin restricted her discussion to Shiller (1979, 1981) and Singleton (1980); therefore it remains to determine whether her criticisms apply to LeRoy and Porter (1981). The criticism just discussed does not apply since LeRoy and Porter made no use of an observable proxy for p^*_t. However, Flavin's first criticism of Shiller, that autocorrelation between successive elements of p^*_t induces downward bias in the variance estimate, does have an analog that applies to LeRoy and Porter. Suppose that dividends follow the first-order autoregression (9). In small samples, the estimated value of the autoregression parameter will be biased toward zero. The first-order term of the bias is *2/n*. This bias induces a corresponding downward small-sample bias in the estimated variance of p^*_t.

However, the bias problem is more complicated than this. As (13) indicates, the variance of p^*_t is a convex function of λ, but then, by Jensen's inequality, the expectation of the sample variance will be greater than the value indicated by (13), with the expected sample value substituted for the population value of each parameter. In sum then, the bias will depend on two effects: λ is estimated with down ward bias, inducing downward bias in $V(p^*_t)$; sample variability in the estimated λ will induce upward bias in $V(p^*_t)$ because of the convexity of the variance expression. The outcome will depend on which effect is stronger.

Thus, about the LeRoy-Porter estimation procedure, Flavin's analysis tells us only what is in any case evident from first principles: any nonlinear estimation procedure will in general be subject to bias, with the sign of the bias impossible to determine *a priori,*

[9]Shea (1989) pointed out that the population variancc of $p^*_{t|T}$ is nonconstant.

2.2 Kleidon

Kleidon's (1986b) criticism focused on Grossman and Shiller's (1981) contention that the smoothness of a time-series plot of p_t^* relative to p_t contradicts the variance-bounds theorems. To Kleidon such a conclusion is completely unwarranted. The variance-bounds inequalities refer to cross sections, not time series. That is, if we could run a large number of replications of the economy, the variance-bounds theorems imply that we would find that

$$\hat{V}(p_t) = \frac{\Sigma_i (p_{it} - \bar{p}_t)^2}{n-1} \leq \hat{V}(p_t^*) = \frac{\Sigma_i (p_{it}^* - \bar{p}_t^*)}{n-1}, \qquad (22)$$

for each t, where i runs over replications of the economy, not time. Unfortunately, history occurs only once, so we have only a single value for i. The time-series inequality

$$\hat{V}(p) = \frac{\Sigma_t (p_t - \bar{p})^2}{n-1} \leq \hat{V}(p^*) = \frac{\Sigma_t (p_t^* - \bar{p}^*)^2}{n-1}, \qquad (23)$$

for a single economy is what the data appear to contradict, but (22) and (23) are very different. Kleidon observed that nothing about the variance-bounds theorems guarantees that (23) will be satisfied, even in large samples. Hence reversal of (23) does not contradict the present-value relation.

To see Kleidon's point, we present two examples. The first is adapted from Kleidon (1986b, pp. 975-976). Suppose that a stock pays a dividend d_T only at some terminal date T. Take d_T to equal $\Sigma_{t=1}^{T} \epsilon_t$, where ϵ_t is an identically and independently distributed random variable with mean zero and variance unity (as above, the mean is suppressed). Assuming no discounting, we have

$$p_t^* = d_T = \sum_{t=1}^{T} \epsilon_t,$$

for all t, so

$$V(p_t^*) = T, \qquad \text{for } t = 1,2,\ldots,T.$$

Actual price is given by

$$p_t = E_t(p_t^*) = \sum^{t} \epsilon_t,$$

implying that

$$V(p_t) = t, \qquad \text{for } t = 1,2,\ldots,T.$$

Since $V(p_t) = t \leq V(p_t^*) = T$, if we look across economies, the variance-bounds inequality is satisfied at each date.

But imagine that only one replication of this experiment is available. The distributions of p_t and p_t^* corresponding to time-series plots of these variables are those conditional on the one realization of d_T. Since $p_t^* = d_T$, for all t, the conditional variance of p_t^* is zero, while p_t evolves randomly from $p_0 = 0$ to $p_T = d_T$. Plainly, p_t^* appears smoother than p_t.[10]

The second example is less transparent, but is nonetheless useful because it aids in the interpretation of Kleidon's simulations. Suppose that dividends follow a geometric random walk,

$$\ln(d_{t+1}) = \ln(d_t) + \epsilon_{t+1},$$

or, equivalently,

$$d_{t+1} = d_t \times \exp(\epsilon_{t+1}), \tag{24}$$

where ϵ_{t+1} is distributed independently as normal with mean μ and variance σ^2 and $d_0 = 1$. Further, if $\mu + \sigma^2/2 = 0$, then $E(d_{t+1} \mid d_t) = d_t$, for all t, so that d_t is a martingale without trend.[11] Now take as a measure of ex post rational price the variable $p_t^*(1)$, defined by

$$p_t^*(1) \triangleq \beta(p_{t+1} + d_{t+1}), \tag{25}$$

so that ex post rational price equals the present value of the position next year. Set p_t in the usual way:

$$p_t = E(p_t^*(1) \mid d_t, d_{t-1}, \ldots). \tag{26}$$

Using the rule of iterated expectations, (25) and (26) imply that

$$p_t = \frac{\beta}{1-\beta} d_t, \tag{27}$$

where we have assumed away bubbles and have used the fact that dividends have an expected growth rate of zero. Therefore p_t is a martingale as well. Because of (26), the variance-bounds theorem implies that p_t has lower volatility than $p_t^*(1)$ for each t. If we take the second moment around zero, we therefore have

$$E[p_t^2] < E[(p_t^*(1))^2],$$

[10] Analytically, we have $V(p_t^* \mid d_T) = 0$, for all t. As to the actual price, its conditional variance is given by

$$V(p_t \mid d_T) = V(p_t) - \frac{[\mathrm{Cov}(p_t, d_T)]^2}{V(d_t)} = t - \frac{t^2}{T}.$$

This parabola has zeros at $t = 0$ and $t = T$. and is positive between the two.

[11] This is so becausc (24) implies that $E(d_{t+1} \mid d_t) = d_t E[\exp(\epsilon_{t+1})]$. The lognormal distribution has the property that $E[\exp(\epsilon_{t+1})] = \exp(\mu + \sigma^2/2) = 1$ if $\mu + \sigma^2/2 = 0$.

for each t, implying that if $\hat{\theta}$ is defined by

$$\hat{\theta} \triangleq \frac{\sum_{t=1}^{T} (p_t^*(1))^2}{T} - \frac{\sum_{t=1}^{T} p_t^2}{T},$$

then

$$E(\hat{\theta}) > 0. \tag{28}$$

The variance-bounds test associated with this inequality-check whether $\hat{\theta}$ is in fact positive-appears straightforward. Neither of Flavin's criticisms applies: by taking second moments around zero instead of the sample mean her first criticism is avoided,[12] while defining ex post rational price so that it is observable avoids her second criticism as Well. Further, dividends have a constant mean by assumption, so it would seem that trend adjustment is not a problem. Yet if the variance-bounds test is conducted on simulated data, $\hat{\theta}$ turns out to be negative in almost every sample which has at least 50 or 100 draws.

Joerding (1986), who originated this example, interpreted it as reflecting a problem with random number generators [see also Joerding (1988), Kleidon and Koski (1992), and Joerding (1992)]. take a different line. Note that, from (25) and (27), we have $p_t^*(1) = p_{t+1}$. Therefore $\hat{\theta}$ takes the simple form

$$\hat{\theta} = \frac{p_{T+1}^2 - p_1^2}{T}.$$

Because p_t is a martingale, the martingale convergence theorem [Billingsley (1979, p. 416)] implies that p_t will-approach a constant on almost every sample path. In view of

$$p_{t+1} = p_t \exp(\epsilon_{t+1})$$

[an Implication of (24) and (27)], the only way p_{t+1} can equal p_t is for the constant to which they converge to be zero. Thus the martingale convergence theorem implies that, in a sufficiently large sample, $T\hat{\theta}$ will be arbitrarily close to $-p_1^2$ with arbitrarily high probability. Even though the variance-bounds inequality is satisfied in the population, its sample counterpart is reversed with arbitrarily high probability in large samples.

The preceding examples give ample indication that a single time. series, no matter how long, may fail to provide reliable estimates of the relevant population variances, but both examples were constructed under the assumption that dividends are nonstationary, a

[12] This device was used by Mankiw, Romer, and Shapiro (1985); see Section 3.

specification explicitly excluded by Shiller, and LeRoy and Porter. We have already seen that if dividends are generated by (9) with $|\lambda| < 1$, then the sample variances $V(p)$ and $V(p^*)$ based on a single time series are asymptotically unbiased estimates of $V(p)$ and $V(p^*)$. However, suppose that $\lambda = 1$, so that dividends are a random walk. Analysis of the geometric version of this case occupies the bulk of Kleidon's article.

Two preliminary points: first, a difference equation such as (9) does not completely characterize dividends, whatever the value of λ. Also needed is an initial condition. In taking $V(p_t)$ and $V(p_t^*)$ as constants, we adopted tacitly the usual convention that the drawing d_0 of dividends at some initial date is from the limiting distribution to which d_t converges. Under this convention, d_t is stationary, implying that its variance is constant. In the random-walk case, however, there. is no limiting distribution. It follows that no matter what distribution d_0 is drawn from, d_t is not stationary. Accordingly, $V(p_t)$ and $V(p_t^*)$, although finite, are functions of time. Thus by taking $\lambda = 1$, Kleidon undertook to analyze the robustness of Shiller's conclusions when the underlying assumption of stationarity is a misspecification. This approach is in sharp contrast to Flavin, who accepted Shiller's stationarity assumption and analyzed small-sample problems.

Second, nonstationarity does not invalidate the variance-bounds theorems. The equation $p_t = E(p_t^*|I_t)$ implies that $V(p_t) \leq V(p_t^*)$, for each t, whether or not dividends are stationary. Rather, it is the assumption that these variances are constant over time, which is adopted in econometric implementation of the variance-bounds theorems, that is violated if stationarity fails;

Now we return to Kleidon's discussion of the Grossman-Shiller contention that the smoothness of p^* relative to p indicates violation of the variance-bounds inequalities. We have seen that if dividends are a random walk, the sample variances of p and p^* over a single time series do not correspond to the (nonconstant) cross-section variances that enter the variance-bounds theorems. Hence, the apparent smoothness of p^* relative to p is uninformative about whether the variance bounds are satisfied.

Sample variances from a single time series, as we have seen, are not good estimators of the corresponding (unconditional) population variances, but that does not mean that they are uninterpretable. We now prove that in small samples

$$E[\hat{V}_n(p_t^m)] > E[\hat{V}_n(p_t^*)] \tag{29}$$

is true if and only if

$$V(p_{t+n}^m|p_t^m) > V(p_{t+n}^*|p_t^*) \tag{30}$$

ante-bounds theorems, which state that the variance of p is less than that of p^* only if the respective variances are conditioned on the same information set. In (33), the conditioning sets are different, so the variance-bounds theorems do not contradict (33). In particular, the variance of p^*_{t+n} conditional on p^*_t has no economic meaning since agents cannot observe p^*_t at t (or, for that matter, at any later date).

Since the variance-bounds theorems do not apply to conditional variances, the inequality (33) must be studied from scratch. This analysis is done easily using the linear autoregression model for dividends. Explicit evaluation of the conditional variances shows that for small n, the inequality (33) is exactly what one should expect. Essentially because the recursion $p^*_t = \beta(d_{t+1} + p^*_{t+1})$ has no error, nearby values of p^* are much more highly correlated than nearby values of p [see LeRoy (1984, pp. 185-186)], leading to conditional variances lower for p^* than for p. If $\lambda = 0.99$, this effect persists for n as high as 50, while if $\lambda = 1$, any finite-sized sample will have the qualitative properties of a small sample.

The foregoing considerations suggest that if dividends follow a random walk, the sample variance of p is likely to exceed the sample variance of p^* for samples of any size, but is the random-walk model consistent with a sample standard deviation of p which exceeds that of p^* by a factor as high as 5.59, the value Shiller reported? To gauge the meaning of the coefficient reported by Shiller, Kleidon ran Monte Carlo studies in which the present-value relation holds by construction and where dividends are generated by a geometric random walk with parameters chosen to give the best fit to Shiller's data. The variance inequality was deemed violated by the sample if the sample variance of p_t exceeded that of p^*_t, and to be "grossly violated" if the ratio of sample variances exceeded 5. Kleidon reported a frequency of violations of about 90 percent (when following Shiller's detrending procedure); the frequency of gross violations varied from 4.6 percent, for a rate of interest of 0.075, to almost 40 percent, for a rate of interest of 0.05.

Shiller (1988) responded that when the rate of interest varies, so does the dividend-price ratio if the dividend growth rate μ stays constant. The combination of the rate of growth μ chosen by Kleidon and a rate of interest of 5 percent implies an implausible value of the dividend-price ratio. Shiller ran experiments similar to those of Kleidon, but specified μ to vary with the rate of interest in order to leave the implied dividend-price ratio at its average historical level. He obtained much lower frequencies of gross violations.

Note that in the debate between Kleidon and Shiller about the appropriate parameter values to assume in Monte Carlo experiments, we have completely lost contact with the variance-bounds theorems.

If dividends follow a random walk (whether arithmetic or geometric), the sample variance bears no relation to any simple unconditional population parameter. Consequently, the robustness of the original variance-bounds theorems is sacrificed. For example, Kleidon and Shiller assumed that agents have no information beyond current dividends. How would their rejection frequencies be affected if this unrealistic assumption were relaxed? In the absence of any underlying theory, it is impossible to say.

Kleidon's paper emphasized that the first-generation empirical tests of the variance-bounds relation depended on the stationarity of the underlying series. The Shiller and LeRoy-Porter papers were clear about this dependence, so the question becomes whether the trend adjustment used in each case in fact produced a stationary series. Shiller's trend removal-take residuals from a constant growth rate path-will produce a stationary series only if the original series is trend stationary. The empirical evidence for the existence of unit roots in dividends-i.e., against the stationarity of dividends-is weaker than for most macroeconomic time series; thus, based on the evidence on dividends alone, the stationarity question remains open. However, in most macroeconomic models, dividends are cointegrated with variables such as GNP, investment, consumption, and so forth. Therefore the relevant evidence is that concerning unit roots for macroeconomic variables generally, not just dividends. The evidence now indicates that most macroeconomic variables do have unit roots. Of course, even if so, it remains open whether the bias induced by nonstationarity is sufficient to explain the variance-bounds violations.

2.3 Summary

The reason Shiller's rejections of the variance-bounds inequalitywere so striking is that the results did not depend on a particular specification of the model for dividends (as noted above, the LeRoy-Porter version did not-share this attractive "model-free" property). Flavin and Kleidon pointed out that, while it is true that the variance-bounds inequality itself is model free, the properties of any econometric test of that inequality can only be investigated conditional on a particular dividends model. They argued that under reasonable specifications of the dividends model, variance-bounds tests will reject with high probability even if the present-value model is true. There can be no doubt that, at a minimum, the critics established that econometric problems with variance-bounds tests are potentially severe. Whether these problems are severe enough to account for the extent of the apparent excess volatility, however, remained controversial.

3. Second-Generation Tests

3.1 west

West (1988a) derived a variance-bounds test that (i) is valid even if dividends are nonstationary, and (ii) does not require a proxy for p_t^*. To understand West's test, recallthe definition of H_t as the information set consisting of current and past dividends. West defined x_{tH} as the expectation of the cum dividend ex post rational price conditional on H_t, and x_{tI} as the corresponding variable for investors' actual information set I_t:

$$x_{tH} \triangleq \sum_{i=0}^{\infty} \beta^i E(d_{t+i} | H_t), \qquad x_{tI} \triangleq \sum_{i=0}^{\infty} \beta^i E(d_{t+i} | I_t).$$

West's result was that under the weak assumption that I_t contains H_t, the innovation variance of x_{tH} exceeds that of x_{tI}:

$$E[(x_{t+1,H} - E[x_{t+1,H} | H_t])^2] \geq E[(x_{t+1,I} - E[x_{t+1,I} | I_t])^2].$$

The meaning of West's inequality becomes apparent when we recast it in more familiar notation. Recalling the definitions of the ex dividends price series p_t and p_t,

$$\hat{p}_t \triangleq \sum_{i=1}^{\infty} \beta^i E(d_{t+i} | H_t), \qquad p_t \triangleq \sum_{i=1}^{\infty} \beta^i E(d_{t+i} | I_t),$$

we have

$$x_{t+1,H} = \hat{p}_{t+1} + d_{t+1} \qquad \text{and} \qquad E(x_{t+1,H} | H_t) = \beta^{-1}\hat{p}_t.$$

Similarly,

$$x_{t+1,I} = p_{t+1} + d_{t+1} \qquad \text{and} \qquad E(x_{t+1,I} | I_t) = \beta^{-1}p_t.$$

Thus the innovations in x_{tH}, and x_{tI} are recognized as just the returns

$$\hat{r}_{t+1} = \hat{p}_{t+1} + d_{t+1} - \beta^{-1}\hat{p}_t \qquad \text{and} \qquad r_{t+1} = p_{t+1} + d_{t+1} - \beta^{-1}p_t,$$

and West's inequality reduces to

$$V(\hat{r}_t) \geq V(r_t). \tag{34}$$

Thus, investors' actual returns must have lower variance than the returns that would obtain if investors' information consisted of H_t.

West's upper bound on return variance is a direct implication of the LeRoy-Porter lower bound on price variance. To see this, observe that the LeRoy-Porter equality (3) holds for any information set. Therefore it holds for H_t implying

$$V(p_t^*) = V(\hat{p}_t) + \frac{\beta^2 V(\hat{r}_t)}{1-\beta^2}. \tag{35}$$

The LeRoy-Porter lower bound (8) on price variance, combined with (3) and (35), yields West's inequality (34).

West argued in favor of his test that, unlike the first-generation variance-bound tests, It is valid whether or not dividends are stationary. Even if dividends are generated by a linear process with a unit root, so that dividends and prices are cointegrated rather than stationary, the population return variances will be constant. Therefore their sample counterparts provide consistent estimates of population values. This argument is correct, but it must be remembered that it applies only for nonstationary dividend processes that are linear. We believe that a more natural treatment of trend is to specify a log-linear, rather than linear, dividend process, so that dividend growth rates are stationary. Although West's variance-bounds test does not apply directly, it can be adapted to the log-linear case; see LeRoy and Parke (1990), discussed below.

3.2 Mankiw, Romer, and Shapiro

Mankiw, Romer, and Shapiro (1985, 1991) (hereafter MRS) claimed to have provided an unbiased volatility test. The derivation is as follows: Let p_t^0 be any variable that can be constructed from information available to agents at t, and let $p_{t|T}^*$, as in Section 2.1, be the observable version of p_t^*. Consider the identity

$$p_{t|T}^* - p_t^0 \equiv (p_{t|T}^* - p_t) + (p_t - p_t^0). \tag{36}$$

Under unbiased forecasting, $p_{t|T}^* - p_t$ must be uncorrelated with any variable observable at t; in particular, with $p_t - p_t^0$. Hence by squaring both sides of (36), taking' expectations, and making use of $E[(p_{t|T}^* - p_t)(p_t - p_t^0)] = 0$, we get

$$S \triangleq E[(p_{t|T}^* - p_t^0)^2] - E[(p_{t|T}^* - p_t)^2] - E[(p_t - p_t^0)^2] = 0. \tag{37}$$

The last equality is valid for any definition of p_t^0, but if p_t^0 is interpreted as a "naive forecast" of $p_{t|T}^*$, then the actual price must outperform p_t^0 as a forecast of $p_{t|T}^*$, which implies (37).

The sample counterpart of (37) is defined in the obvious way:

$$\hat{S} \triangleq \frac{1}{T}\left\{\sum_{t=1}^{T} (p_{t|T}^* - p_t^0)^2 - \sum_{t=1}^{T} (p_{t|T}^* - p_t)^2 - \sum_{t+1}^{T} (p_t - p_t^0)^2\right\}.$$

MRS set p_t^0 as a cynstant multiple of dividends in the preceding year. They found that S was negative, indicating excess volatility.

MRS claimed that because this volatility test uses noncentral rather than central variances, it is unbiased.[15] To evaluate this claim, we first point out a redefinition of terms that crept into the variance-bounds literature with Flavin's paper. By "test," econometricians ordinarily mean something very precise. Besides specifying a test statistic, it is necessary to identify a rejection region such that, if the null hypothesis is true, it will be rejected with preassigned probability a. A test is biased toward rejection if the probability of rejection exceeds α.

In the variance-bounds literature, however, the terms "test" and "unbiased test" have been used more loosely. Because of the difficulty of evaluating small-sample distributions, many have dispensed with any formal attempt to specify a rejection region as defined above. The variance-bounds inequalities are then rejected if the sample variance of p_t is much higher than that of p^*_t, with no attempt to define "much." In the variance-bounds literature, a test 'is said to be biased toward rejection if the expected value of the test statistic is in the direction of excess volatility, relative to the corresponding population parameter. Because this property concerns only the first moment of the distribution of the test statisic, it is usually easy to investigate. For example, the equality $E(S) = S = 0$ is the basis for MRS's claim to have provided an unbiased test. However, the result that S gives an unbiased estimate of S does not establish that the test is unbiased in the usual sense.

Given the difficulty in evaluating type I error probabilities, the redefinitions of the terms "test" and "unbiased test" that have taken place are understandable. The two definitions are related: in showing small-sample bias (as redefined), Flavin created a strong presumption that the apparently dramatic rejections of the variance-bounds inequalities that Shiller reported are suspect, but when MRS went on to report the results of what they called an unbiased test, it is important to note that by this they meant only that the expectation of their test statistic equals the corresponding population parameter.

Shea (1989) pointed out two major problems with MRS's tests, both attributable to the use of $p^*_{t|T}$ in place of p^*_t. First, the outcome of the tests is very sensitive to the choice of terminal date. Second, because of the nonstationarity induced by the dependence of both population parameters and statistics on $p^*_{t|T}$ (which is nonstationary even if p^*_t is stationary), there is no prospect of using asymptotic theory to derive confidence intervals for the tests. To avoid both problems, Shea suggested using a rolling terminal date. Under Shea's

[15] Flavin had already noted that noncentral variant have this property and had reported the results of such tests. She credited Richard Porter with the insight that use of noncentral variances provides a way around her critism of Shiller's tests (1983. p. 950).

recommended procedure, for each time period the same value of τ is used to calculate the ex post rational price

$$p^*_{t|\tau} \triangleq \sum_{i=1}^{\tau} \beta^i d_{t+i} + \beta^\tau p_{t+\tau},$$

instead of defining τ as T - *t*. However, the estimates of $p^*_{t|\tau}$ calculated in this way are not equal to the expectation of p^*_t conditional on the entire sample, lessening the appeal of Shea's procedure. Also, Shea's procedure is wasteful of data: with sample size *T*, a rolling terminal date τ allows construction of the price series only until $t = T - \tau$ since $p^*_{t|\tau}$ requires an observation of $p_{t+\tau}$. A larger value of τ reduces the observation error in the series $p^*_{t|\tau}$, but also reduces its length. In any case, Shea's results were much more favorable to the variance-bounds theorems than MRS's.

As noted above, in their 1985 paper, MRS made no attempt to determine whether their rejection of the present-value model was statistically significant-they concluded only that their point estimates suggested excess volatility. However, in their 1991 paper, MRS calculated asymptotic standard errors, using the rolling horizon to assure stationarity. They concluded that the excess volatility was of moderate but not overwhelming statistical significance. This was exactly the conclusion LeRoy and Porterhad reached, also based on asymptotic standard errors.

Marsh and Merton (1986) and Merton (1987) observed that dividend smoothing by management could' bias variance-bounds tests in general, and MRS's test in particular, toward. rejection. If dividends are slow to reflect changes in underlying profitability, measured dividend volatility could give an impression that fundamentals had remained stable-even when the opposite was the case. To illustrate theproblem, Merton (1987) analyzed an example in which dividends were assumed to be zero within the sample and showed that econometric problems with the MRS test would result. Merton's example is a hybrid of models analyzed by Flavin and Kleidon. The nonstationarity of prices implied by the zero-dividend assumption results in econometric problems, as Kleidon showed; the substitution of the observable $p^*_{t|T}$ for the unobservable p^*_t induces small-sample bias, as Flavin showed. In fact, the latter problem is magnified in the context of Merton's example as a consequence of his assumption that within-sample dividends are zero. This observation formed the basis for MRS's (1991) response to Merton's criticism.' They noted that, empirically, end-of-sample price is a small contributor to within-sample price relative to within-sample dividends at most dates of a hundred-year sample. Therefore, they argued, the econometric problem Merton pointed out is of little practical importance.

Merton's criticism and MRS's reply failed to distinguish between two separate phenomena: low dividends and smoothed dividends. It is true that firms retain a large fraction of their earnings, implying that capital gains typically exceed dividends, but that does not necessarily cause problems for model-based variance-bounds tests (since these estimate the variance of ex post rational price directly from dividends, rather than relying on some observable version of p_t^*). For example, if dividends follow a geometric random walk, variance-bounds tests can be constructed by estimating the mean and variance of the dividend growth rate and comparing the implied volatilityof ex post rational price with that of actual price [see LeRoy and Parke (1990), discussed below]. Nothing about Merton's example suggests that these parameters are more difficult to estimate when the mean dividend growth rate is high (as will occur when corporations retain most of their earnings) than when it is low. Thus low dividend payout rates, even though they imply that stock prices depend mostly on out-of-sample dividends, do not necessarily cause problems for variance-bounds tests, Smoothed dividends, however, are another story. If dividend growth rates have important low-frequency components, small-sample problems are exacerbated. MRS's reply did not address this aspect of Merton's criticism.

3.3 Scott, and Durlauf and Hall

Scott (1985) observed that nothing about the hypothesized-equality between p_t and $E(p_t^*|I_t)$ requires that this equality be tested by comparing variances. Under the null hypothesis, the error e_t in

$$p_t^* = a + bp_t + e_t$$

is uncorrelated with the explanatory variable, so the present value model can be tested equally well by determining whether a equals zero and b equals unity in an ordinary least squares regression. Scott argued that such regression tests of the present-value model are essentially free of the econometric problems associated with volatility tests. He found empirically that b was near zero rather than unity, implying that p_t^* and p_t are virtually uncorrelated, conflicting with the present-value model.

Recent work by Durlauf and Hall (1989) clarified the. relation between Scott's regression test and the volatility tests of the present-value relation. Continuing along the line initiated by Scott, Durlauf and Hall pointed out that it is possible to translate the restriction on parameters implied by volatility tests into a corresponding coefficient restriction in regression tests. To achieve this translation, Durlauf and Hall modeled stock prices p_t *as* the sum of q_t, the expected value of discounted dividends, and a noise variable s_t, which is unrestricted.

As usual, p_t^* is written as q_t plus an orthogonal forecast error f_t. We have

$$p_t = q_t + s_t, \qquad p_t^* = q_t + f_t, \tag{38}$$

where the left-hand side variablesare observed by the econometrician and the right-hand side variables are unobserved. The present-value model corresponds to $s_t = 0$. Now consider the regression

$$p_t = \theta(p_t - p_t^*) + u_t.$$

We wish to obtain the restriction on θ implied by $V(p) \leq V(p^*)$. To do *so*, write $\mathrm{Cov}(p_t, p_t - p_t^*)/V(p_t - p_t^*)$ and use (38) to eliminate P_t and p_t^*:

$$\begin{aligned}\theta &= \frac{\mathrm{Cov}(q_t + s_t, s_t - f_t)}{V(s_t - f_t)} \\ &= \frac{\mathrm{Cov}(q_t, s_t) + V(s_t) - \mathrm{Cov}(s_t, f_t)}{V(s_t) - 2\mathrm{Cov}(s_t, f_t) + V(f_t)}\end{aligned} \tag{39}$$

[using the orthogonality condition $\mathrm{Cov}(q_t, f_t) = 0$]. Similarly, express the variance bounds inequality $V(p_t) \leq V(p_t^*)$ in terms of the moments of s_t, q_t, and f_t:

$$V(q_t) + 2\,\mathrm{Cov}(q_t, s_t) + V(s_t) \leq V(q_t) + V(f_t). \tag{40}$$

Combining (39) and (40), Durlauf and Hall obtained $\theta \leq \frac{1}{2}$.

Thus, rejection of $V(p) \leq V(p^*)$ is equivalent to rejecting $\theta \leq \frac{1}{2}$, but the present-value model should be rejected for any θ that differs significantly from zero, not just for values of θ that significantly exceed $\frac{1}{2}$. Durlauf and Hall argued that regression-based tests such as that of Scott, which reject for any nonzero value of θ, are likely to be superior to volatility tests, which do not detect small amounts of model noise.[16]

This argument is incorrect. The reason is that so far we have been vague about the choice of rejection region. If in a model-based variance-bounds test the rejection region is computed via simulation, then-the rejection region associated with the chosen probability of type I error depends on the variance of the forecast error. In the Durlauf-Hall regression context, the same statistic, θ, is involved in both the (bounds) test of the inequality $\theta \leq \frac{1}{2}$ and the (orthogonality) test of the equality $\theta = 0$. 'Therefore type I error can be set equal in the two cases only if both tests have the same rejection region, but then the probability of type II error will also be equal under any

[16] LeRoy and Porter (1981, p. 561). Frankel and Stock (1987), and hoot (1989). reasoning similarly, also concluded that bounds tests have lower power than orthogonality tests.

alternative hypothesis so that, contrary to the Durlauf-Hall argument, in the regression context orthogonal&y tests do not have greater power than bounds tests.

The main purpose of the Durlauf-Hall paper was not to compare bounds and orthogonal&y tests of stock price volatility, but to measure the variance of model noise as a contributor to the total variance of stock prices. The fact that noise cannot be assumed to be orthogonal to any of the variables causes ambiguity in the measurement of noise variance. However, Durlauf and Hall showed that there exist lower bounds on the amount of model noise consistent with the data. Durlauf and Hall showed that at least 84 percent of the variance of prices can be attributed to noise under the estimator associated with the bounds test $V(p) \leq V(p^*)$.[17] Under the noise variance estimator associated with the more stringent orthogonality test, noise variance is essentially equal to that of stock prices. Durlauf and Hall concluded from these results that "movements in expected discounted dividends, with a constant discount rate, have essentially nothing to do with the actual movements of the stock market" (p. **17).**

3.4 Campbell and Shiller

In three recent papers, Campbell and Shiller (1987, 1988a, 1988b) introduced material not found in Shiller's early papers on asset-price volatility. Campbell and Shiller noted that if the present-value model is true, (i) an optimal prediction of the present value of future expected dividends can be formed using current price alone and (ii) this optimal prediction coincides with current price. It follows that the present-value model implies testable restrictions of the coefficients of a bivariate vector autoregression of stock prices and dividends. This insight was not new to the variance-bounds literature (the LeRoy-Porter volatility test consisted exactly of tests of restrictions on the coefficients of a bivariate autoregression).[18] However, in implementing their tests, Campbell and Shiller (1988a) introduced a method for trend correction that, although not perfect, is superior to anything that went before. Campbell and Shiller assumed that dividends and whatever other variables predict dividends form a multivariate log

[17] Apparently, Durlauf and Hall did not correct for trend in any way. Therefore, given that the foreast error is a martingale (facing toward the past), if the price process has a unit toot, then spurious correction will result, implying upward bias in the Durlauf-Hall estimate of noise variance. We are indebted in Douglas Steigerwald for pointing this out.

[18] In 1987, Campbell and Shiller observed that "there ate several [procedures for testing the present-value relation] In the literature: these include single-equation regression tests, tests of cross-equadon restrictions on a vector autoregression (VAR), and variance bounds tests" (p. 1063). In taking the latter two categories to be distinct, Campbell and Shiller appeared to be unaware that the Ida of testing the volatility implications of the present-value relation by investigating the validity of cross-equation restrictions in a bivariate autoregression had already been introduced into the vatlance-bounds literature.

linear model. To reconcile the log-linearityof the dividend model with the linearity of the present-value relation, Campbell and Shiller log-linearized the expression defining the rate of return (the opposite procedure, linearizing the dividend model, would introduce heteroskedasticity). Comparison of the actual rate of return with its log-linearized counterpart-the two are correlated almost perfectly-allowed Campbell and Shiller to argue that the error introduced by the log-linearization is negligible. The present-value model that results from iterating the log-linearized version of the definition of the rate of return expresses the log price-dividend ratio as the present value of the discounted expected dividend growth rates. All these variables are essentially free of trend, implying that the econometric problems attending the first-generation volatility tests by reason of the nonstationarity of the underlying series are avoided.

Campbell and Shiller reported the results of a variety of tests of the equality of the log price-dividend ratio and the present value of future dividend growth rates, and for the most part found evidence of significant violation. Most relevant for the variance-bounds question, for example, Campbell and Shiller compared the standard deviation of the log actual price-dividend ratio with the standard deviation of the present value of future expected dividend growth rates. If the present value model is correct, it must be the case that, since these variables are the same, their standard deviations are equal. In fact, Campbell-Shiller found that the standard deviation of the actual price-dividend ratio was twice that of its theoretical counterpart, indicating significant excess volatility.

In 1988, Campbell and Shiller (1988b) added corporate earnings to the price-dividend autoregression. They found that earnings is a strong predictor of dividend growth even conditional on the current log price-dividend ratio. This finding contradicts the present-value model, according to which current price is a sufficient statistic for future dividend growth. This article. also contains a valuable discussion of the relation between the variance-bounds tests and the return autocorrelation tests conducted by Fama and French (1988), Poterba and Summers (1988), and others. Recall that LeRoy and Porter had based their variance-bounds test on the fact that the forecast error for ex post rational price equals a weighted average of future total returns, so rejection of the variance-bounds inequality implies that returns are autocorrelated. Consequently, in the language of Campbell and Shiller, “excess volatility and predictability of multiperiod returns are not two phenomena, but one” (p. 663).

Unfortunately, direct comparison between the first-generation variance-bounds tests and the return autocorrelation results is impossible because the Former tests imply autocorrelatedness of total return, whereas the latter tests imply autocorrelatedness of the rate of return (i.e., total return per dollar of stockvalue). However, under the Campbell-Shiller linearization, the analog of total return is exactly the rate of return. Thus, since the same return measure is involved in both variance-bounds and return autocorrelation tests, one can go beyond the mere observation that rejection of the variance-bounds inequality implies returns are autocorrelated. In addition, one can translate any pattern of autocorrelation of rates of return into the implied correlation between the log actual price-dividend ratio and the forecast error for log ex post rational price-dividend ratio. Further, rejection of the orthogonality of the log ex post rational price-dividend ratio and the log actual price-dividend ratio implies exactly that a weighted average of future rates of return, with weights that decline geometrically according to a factor related to the average price-dividend ratio, is forecastable. Given the numerical value of this weighting factor, it turns out that the variance-bounds violations correspond to the finding that long-term (five- and ten-year) rates of return are highly autocorrelated. This was exactly the finding of Fama and French and Poterba and Summers.

3.5 LeRoy and Parke, and Cochrane

LeRoy and Parke (1990), in work initially conducted independently of the Campbell-Shiller research discussed above, proposed correcting for the trend in stock prices by dividing prices by dividends. They assumed that dividends follow a geometric random walk, their purpose being to derive variance-bounds tests that would be valid in the setting used by Kleidon to criticize Shiller. LeRoy and Parke began by reporting a model-free bounds test of the volatility of the price-dividend ratio along the lines of Shiller (1981), and also a model-based bounds test as developed by LeRoy and Porter. Taking the rejection-region to be defined by a negative value of the volatility statistic $\hat{V}(p^*) - \hat{V}(p)$, they found that the model-free bounds test rejected the model, whereas the model-based test accepted it. This discrepancy reflected primarily the radically different estimates of ex post rational price volatility generated under the two procedures: the volatility estimate produced by the model-free. procedure [using $p^*_{t|T}$, the observable version of ex post rational price obtained from the recursion (15) and the terminal condition $p^*_{T|T} = p_T$] was less than one-quarter that produced by the model-based procedure (estimating the discount factor and the mean and variance of the dividend growth rate and entering these as arguments in the expression for the variance of ex post rational price).

The obvious explanation for the disparity between model-free and model-based estimates of ex post rational price volatility is estimation bias. Flavin's arguments established the strong presumption that the model-free estimates are biased downward. The model-based estimates are also subject to bias: even though the mean and variance of the dividend growth rate and discount factor can be estimated without bias, sampling variance in the estimates of these parameters will generally induce bias in the estimated variance of ex post rational price because of the nonlinearity of the variance expression.

This differential effect of estimation bias suggests that type I error is much higher under model-free tests than under model-based tests. This reflects the arbitrary specification that the rejection region consists of negative values of $\hat{V}(p^*) - \hat{V}(p)$ regardless of sampling error. A better procedure would be to use the Monte Carlo simulations to specify a rejection region that fixes the estimated probability of type I error at some constant like 10 percent. LeRoy and Parke found that the 10 percent rejection region (based on the Monte Carlo experiments) depended critically on how much information agents are assumed to have, which is not restricted under the null hypothesis. The first conclusion of LeRoy and Parke, therefore, was that the inequality test of the variance-bounds relation, being subject to a "nuisance parameter" problem, is uninformative about the null hypothesis. This is true of both the model-free and model-based versions.

LeRoy and Parke went on to consider orthogonality tests. The orthogonality test (7), devised by LeRoy and Porter, takes the form of an equality, not an inequality. Because under the null hypothesis this equality is valid. for any specification of agents' information, there is no nuisance parameter problem (at least with regard to population parameters; there remains the possibility that sampling distributions depend on agents' information). The problem with (7) is that its derivation depends on the assumption that dividends are generated by a linear process, which is inconsistent with the log-linear geometric random walk. To get around this problem, LeRoy and Parke derived an analog to (7) that is valid under log-linear dividend processes. The associated null hypothesis was significantly rejected for any specification of agents' information.

Finally, LeRoy and Parke considered West's test. The Campbell-Shiller log-linearization implies that West's inequality takes a particularly simple form under the geometric random walk: it says that the variance of the rate of return on stock is bounded above by the variance of the dividend growth rate. As an inequality test, West's test is in principle subject to the same nuisance parameter problem as the price variance inequality test. In practice, however, it turned out that there was no nuisance parameter problem. The Monte Carlo tests showed that, as with the orthogonality test, West's inequality was significantly rejected for any specification of agents' information.[19]

The LeRoy-Parke results depend on two restrictive assumptions: that dividends follow a geometric random walk and that the discount rates are constant. Neither assumption can be justified except as a crude approximation. However, the problem with more complex and realistic models for dividends and discount factors is that under such specifications the rational expectations assumption becomes less plausible. For example, it is easy to verify that dividend growth rates have nonzero autocorrelations; contrary to the LeRoy-Parke specification, but it strains one's credulity to assume that agents know a sequence of autocorrelation coefficients for dividend growth rates and price stock taking account of these deviations from randomness.

Perhaps it is a matter of taste whether one prefers simple models that are not descriptively. accurate or more complex models that, because of their complexity, put the rational expectations assumption to harder use. In any case, Cochrane (1990) derived a version of the LeRoy-Parke bound on the variance of the pricedividend ratio that is consistent with time-varying discount rates and autocorrelated dividend growth rates. To evaluate the bound, Cochrane used a linear approximation similar (but not identical) to that of Campbell and Shiller. Cochrane found that the variance bound was satisfied. This result is analogous to the LeRoy-Parke finding that the model-based point estimate of the variance of the price-dividend ratio was less than its upper bound. As discussed above, LeRoy and Parke concluded that in the absence of an attempt to control the probability of type I error, this result cannot be construed as favoring the present-value relation. The same observation applies to Cochrane's finding.

Cochrane also estimated decompositions of the variance of the price-dividend ratio into the sum of expected dividend growth rates and discount factors along the lines of Campbell and Shiller. Finally, he derived bounds on the mean and standard deviation of discount factors, as in Hansen and Jagannathan (1991). This material, being related only indirectly to variance bounds, is not discussed here.

3.6 Summary

The critics had established that, as attractive as Shiller's model-free variance-bounds test was by reason of its generality, it had the defect that its econometric properties cannot be investigated. Further, under dividend models that were (represented as) reasonable, these prop-

[19] See LeRoy (1990) for a nontechnical exposition and for a plot of the fate of return on stock against the dividend growth rate.

erties are very unsatisfactory. The emphasis in the second-generation variance-bounds tests was on developing tests that had acceptable econometric properties under realistic dividend models. In this they largely succeeded. With the exception of Mankiw, Romer, and Shapiro (1991) and Cochrane (1990), the second-generation tests found statistically significant excess volatility. However, it is important to recall that any model-based test is necessarily a test of a joint hypothesis that includes an assumption about the process generating dividends. Consequently, there remains the possibility that the appearance of excess price volatility may be caused by misspecification of the dividend model.

4. Possible Explanations

In this section, we review the suggested explanations for excess volatility of asset prices. First, investors may overreact to dividend news. That is, market participants may not be perfectly rational and have rational expectations. According to this view, stock prices are moved by fads [Shiller (1984), DeBondt and Thaler (1985)]; or, as Keynes described it, the stock market is a beauty contest in which one wins by "anticipating what average opinion expects the average opinion to be" (1936, p. 156). DeLong et al. (1988) showed that if "noise traders" have irrationally optimistic expectations, there is an equilibrium in which stock prices are more volatile than under the present-value model; Moreover, this equilibrium is stable in the sense that noise traders do not necessarily suffer losses.

Within the paradigm of neoclassical economics, rationality is part of the inner core, insulated from falsification. Explanations for price volatility that involve irrationality are therefore inadmissible. Before jettisoning the neoclassical paradigm to explain one (admittedly important) phenomenon, one should consider explanations that do not rely on irrationality.

If stock prices have greater volatility than expected values of discounted future dividends, the reason may be rational speculative bubbles.[20] The theory and evidence on bubbles was reviewed in West (1988b) and Flood, Hodrick, and Kaplan (1986). If the price of a security contains a deterministic bubble, then the value of-the bubble

[20]See Gilles and LeRoy (1990) for a discussion of rational speculative bubbles within a context of general equilibrium. Flood and Hodrick (1986) observed that model-free variance-bounds tests include bubbles in the null hypothesis because of the way the observable version of ex post rational price is constructed. Therefore a finding of excess volatility in a non-model-based test cannot be attributed to bubbles. Flood, Hodrick, and Kaplan (1986) noted that the model-based tests of LeRoy and Porter did not use the observable version of a post rational price, implying that in this case the presence of bubbles could cause rejection of the variance-bounds inequality,

will grow at the rate of interest, inducing nonstationarity of the price even when dividends are stationary. The data on price-dividend ratios, however, show no evidence of the implied nonstationarity [Diba and Grossman (1988)].

Stochastic bubbles are more interesting. They grow, then collapse. A stochastic bubble, once having burst, cannot be reborn;. if there exists such a bubble, it must have been present from the asset's inception. However, there. may exist many stochastic bubbles on the same price, and they may start so small that they do not noticeably affect the price of the asset for a long period before taking off and then bursting. To the observer, the price will exhibit swings. If bubble bursts are independent of dividends, stock prices are more volatile than the present-value relation predicts. If bursts are correlated negatively with dividends, still greater excess price volatility results and markets appear to overreact to dividend news.

West (1987) at first favored the bubble explanation of stock-price volatility., Upon reviewing the evidence in West (1988b), however, he reversed his conclusion. One of the reasons he gave for disqualifying rational bubbles Is that they cannot be the source of excess volatility In bond prices. Bubbles, whether deterministic or stochastic, can only occur when the horizon is infinite, at least in discrete-time models. With a finite horizon, a backward induction annihilates any bubble at its inception. Shiller (1979) and Singleton (1980) reported evidence of excess volatility of bond prices. Because bonds have finite maturity, such excess volatility cannot be due to bubbles. It seems likely, West argued, that stock prices show too much variability for the same reason that bond prices do; if so, the culprit is not bubbles. Against this argument, however, is the new evidence [Campbell and Shiller (1987)] that the Shiller-Singleton finding of excess volatility in the bond market is due to incorrect trend adjustment. If so, the apparent excess volatility of stock prices, even under correct trend adjustment, favors the bubble explanation:

The present-value relation as we defined it assumes that the rate of discount is constant (see note 1). Since this assumption does not necessarily, hold, it is natural to study the effect of discount rate variability on asset-price volatility. To see the connection, assume that there is a single consumption good and that the representative consumer maximizes $E(\Sigma_j \beta^j u(c_{t+j}) \mid I_t)$ with u of the constant relative risk aversion class $u(c) = c^{1-\rho}/(1-\rho)$. Let r_{it} denote the gross one-period rate of return (1 plus the rate of return) on asset i, and let m_t be the marginal rate of substitution of consumption at $t+1$ for consumption at t [i.e.; $m_t = \beta u'(c_{t+1})/u'(c_t)$]. Assuming interior maxima, the following relation will hold [Grossman and Shiller (1981)]:

$$1 = E(m_t \times r_{it} \mid I_t). \tag{41}$$

If the consumer is risk neutral ($\rho = 0$), then (41) yields the constant expected rate of return (the same for all securities) $E(r_{it}) = 1/\beta$, and hence we get the present-value relation that the data reject. In this model, excess volatility (relative to that predicted by the present-value model) implies that the consumer is not risk neutral.

Grossman and Shiller (1981), LeRoy and LaCivita (1981), and Michener (1982) worked out the implications of (41) for the volatility of asset prices. All found that risk aversion implies greater price volatility than in the present-value model. These results seem encouraging, but for the utility index ln(c) to produce violation of the variance. bounds of the observed magnitude; the aggregate consumption series must be much more volatile than it actually is. The very smooth actual post WWII consumption series implies an implausibly high estimated value of the coefficient of relative risk aversion ρ [Grossman and Shiller (1981)].

The Grossman-Shiller finding of a high value of ρ is consistent with that of Mehra and Prescott (1985), who based their approach on the average covariance between the marginal rate of intertemporal substitution and stock returns. They considered the implications of (41) for the equity premium, defined as the difference between the average rate of return on a diversified portfolio of stocks and the risk-free rate. Mehra and Prescott found that the average real annual yield on the Standard and Poors 500 Index was 7 percent, while that on short-term debt was less than one percent. Let the latter of these yields stand in for the risk-free rate r^* and note that (41) implies

$$r_t^* = 1/E(m_t \mid I_t). \tag{42}$$

We can then rewrite (41) as

$$\frac{E(r_t \mid I_t) - r_t^*}{r_t^*} = -\mathrm{Cov}(m_t, r_t \mid I_t),$$

where r_t is the gross return on the Standard and Poors 500 Index. Mehra and Prescott constructed an artificial economy parameterized so that the growth rate of consumption is a stationary random process with the same mean, variance, and serial correlation as those observed in the U.S. economy. They concluded that an elasticity of substitution between the year t and year $t + 1$ consumption good sufficiently low to yield the 6 percent average equity premium also yields real interest rates far in excess of those observed. Because the elasticity of intertemporal substitution is the reciprocal of ρ, an equity premium equal to six times the risk-free rate implies a high value of ρ, precisely the Grossman-Shiller finding. The main conclusion of Mehra and Prescott was that, ‘even if one were to accept a high value of ρ as plausible, the model (41) is not saved because (42) then implies too high a value for the risk-free rate.

Mehra and Prescott calibrated an artificial economy, rather than using standard econometric techniques. It is, however, possible to modify methods used in the present-value model to test the specific model with time-varying discount rates [Campbell and Shiller (1988a)]. Singleton (1987) reviewed the literature on consumption-based asset-pricing models and reported on the results of a large number of studies besides those noted here. He concluded from these results that the U.S. aggregate data reject the model. This conclusion is robust to changes in the specification of the utility function and in the number of goods assumed to be present.

Progress on these issues will almost certainly require introducing either frictions or non-von Neumann-Morgenstern preferences. Mehra and Prescott and others suggested that the equity premium puzzle points to the importance of frictions. Bewley (1982), following up on this suggestion, showed that if consumers are liquidity constrained, or can only imperfectly insure wealth fluctuations, then individual consumption streams do not fluctuate in sympathy with each other, so the aggregate consumption series will be smoother than the individual series. A definitive test of Bewley's model would require panel data on consumption, asset holdings, and income of individuals, data that are not readily available.

The main problem for the representative-agent expected-utility model is that the smoothness of the aggregate consumption series suggests a low elasticity of intertemporal substitution. A variety of other evidence suggests that the coefficient of relative risk aversion is also low (or moderate), but within the expected-utility, time-separable framework, the intertemporal elasticity of substitution is the reciprocal of the coefficient of relative risk aversion, so we cannot have both. It is possible that uncoupling risk aversion and intertemporal substitution holds the key to the volatility and equity-premium puzzles. Relaxing time separability while preserving expected utility gives enough flexibility to distinguish behavior due to temporal substitution from that due to risk aversion. However, as explained in Epstein and Zin (1989), nonseparable utility leads an expected-utility maximizer to choose consumption plans that are dynamically inconsistent: a plan chosen as optimal in one period becomes suboptimal later. Preferences that do not vary in this fashion and also induce behavior toward risk that is divorced from intertemporal choice cannot satisfy von Neumann-Morgenstern axioms; a consumer with such preferences cannot be indifferent to the timing of the resolution of uncertainty [Kreps and Porteus (1978)], as expected utility implies.

However, Kreps-Porteus or still more general preferences could reconcile the low degree of intertemporal substitution indicated by aggregate consumption with. the moderate degree of risk aversion that economists consider reasonable. Epstein and Zin (1989) showed that in such a setting, the risk premium on a particular asset depends on the correlation of the asset's return both with the market portfolio-as in the capital asset pricing model – and with consumption-as in (41). Epstein and Zin (1991) conducted an empirical study with mildly encouraging results: their stock market and aggregate consumption data rejected a von Neumann-Morgenstern specification of preferences, but could not reject a Keps-Porteus specification. Weil (1989), however, showed that Kreps-Porteus preferences did not solve the risk-premium puzzle, but instead uncover two puzzles that Mehra and Prescott could not distinguish in the expected-utility framework. In addition to the risk-premium puzzle (the risk premium is too high, in view of the evidence on risk aversion), there is a risk-free rate puzzle (the risk-free rate is too low, in view of the evidence on intertemporal substitution). To explain these puzzles, Weil concluded, a proper specification of preferences is probably less important than market frictions.

5. Conclusion

In its 15-year history, the variance-bounds literature has evolved on a circuitous path. Shiller's first-generation model-free tests provided point estimates of price variances that suggested excess volatility, but the question of statistical significance remained open. This was necessarily the case because, in the absence of an assumed model for dividends, there was no way to investigate the econometric properties of the volatility, statistics. The critics remedied this deficiency by supplying models for dividends and deriving the econometric properties of the variance-bounds statistics under the joint assumption that the present-value model is correct and that the dividends equation is well specified. They identified various sources of bias in the model-free variance-bounds tests, suggesting that the apparent excess volatility of asset prices was spurious. The second-generation tests took the obvious next step: expand the null hypothesis to include a model for dividends and then specify a variance-bounds test that performs well if the null hypothesis is true. The model-based tests that resulted-for example, those of Campbell and Shiller and LeRoy and Parke-have much In common with the first-generation model-based tests.

If in some respects the variance-bounds literature evolved toward a point near to where it started; in other respects it evolved into new territory. For example, adequate solutions to the problem of detrending price data, which had been the downfall of both the first-generation papers, were offered in the second-generation papers. Also, the second-generation papers brought to the forefront several critical distinctions that, although present in the first-generation papers, were not emphasized adequately or developed there, Besides the distinction between model-based and model-free tests, one thinks of the distinction between bounds and orthogonality tests.

The evidence is in: asset prices are more volatile than is implied by the present-value equation. Because of the vast improvement in our understanding of the econometric issues attending the variance-bounds tests, there is no longer any room for reasonable doubt about the statistical significance of the excess volatility. Several possible explanations for the excess volatility were reviewed in the preceding section. The authors of this paper are not in complete agreement about how promising these lines of research are. The first author expects that taking adequate account of frictions will go a long way toward resolving the excess volatility, particularly under parameterizations of preferences that are more general than those considered so far in the variance-bounds literature. The second author hopes this is correct, but expects that when the explanation of excess volatility is found, the critical ideas will lie farther from the neoclassical paradigm.

References

Bewley, T. P.. 1982. "Thoughts on Tests of the Intertemporal Asset Pricing Model," working paper, Northwestern University, July.

Billingsy, P., 1979, *Probability and Measure* (2nd ed.), Wiley; New York.

Campbell, J. Y., and R. J. Shiller, 1987, "Cointegration and Tests of Present Value Model," *Journal of Political Economy*, 95, 1062-1088.

Campbell, J. Y.. and R. J. Shiller. 1988a. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies* 1, 195-228.

Campbell, J. Y., and R. J. Shiller, 1988b, "Stock Price, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661-676.

Cochrane. J. C.. 1990. "Explaining the Variance of Price-Dividend Ratios," working paper, University of Chicago.

DeBondt, W. F. M., and R. Thaler, 1985, "Does the Stock Market Overreact?" *Journal of Finance*, 40, 793-608.

DeLong, J. B., et al., 1988, "Noise Trader Risk in Financial Markets," Discussion Paper 1416, Harvard Institute of Economic Research, December.

Diba, B., and H. Grossman, 1988, "Explosive Rational Bubbles in Stock Prices," *American Economic Review*, 78, 520-530.

Durlauf, S. N., and R. E. Hall, 1989, "Measuring Noise in Stock Prices," working paper, Stanford University.

Epstein, L G., and S. E. Zin, 1989, "Substitution, Risk Aversion, and Temporal Behavior of Con sumption and Asset Returns A Theoretical Primework," *Econometrica,* 57,937-969.

Epstein, L G., and S. E. Zin, 1991, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy,* 99, 263-286.

Fama, E. F., and K. R. French, 1988. "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy,* 96, 246-273.

Flavin, M. A., 1983, "Excess Volatility in the Financial Markets: A Reassessment of the Empircal Evidence," *Journal of Political Economy,* 91, 929-956.

Flood, R. P., and R. J. Hodrick, 1986. "Asset Price Volatility, Bubbles, and Process Switching," *Journal of Finance,* 41, 831-842.

Flood, R. P., R. J. Hodrick, and P. Kaplan, 1986, "'An Evaluation of Recent Evidence on Stock Market Bubbles," working paper, Northwestern University.

Frankel, J. A, and J. Stock, 1987, "Regression vs. Volatility Tests of Foreign Exchange Markets," *Journal of International Money and Finance,* 6, 49-66.

Froot, K. A., 1989, 'Tests of Excess Forecast Volatility in the Foreign Exchange and Stock Markets," working paper, MIT.

Gilles, C., and S. P. LeRoy, 1990, "Bubbles and Charges," working paper, University of California, Santa Barbara; forthcoming in *International Economic Review.*

Grossman. S. J., and R. J. Shiller, 1981. "The Determinants of the Variability of Stock Market Prices," *American Economic Review,* 71, 222-227

Hansen. L P., and R. Jagannathan, 1991, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy,* 99, 225-262.

Joerding, W., 1986, "Variance Bounds Tests and Simulated Stock Prices," working paper, Wash. ington State University.

Joerding, W., 1988, "Are Stock Prices Excessively Sensitive to Current Information?" *Journal of Economic Behaviour and Organization,* 9, 71-85.

Joerding, *W.*, 1992, "Are Stock Prices Excessively Sensitive to Current Information?: Reply," *Journal of Economic Behaviour and Organization,* forthcoming.

Keynes, J.M., 1936, The General Theory of Employment, Interest, and Money, Harcourt, Brace &

Kleidon, A. W., 1986a, "Bias in Small Sample Tests of Stock Price Rationality," *Journal of Business,* 59, 237-261.

Kleidon, A. W.. 1986b. "Variance Bounds Tests and Stock Price Valuation Models," *Journal of Political Economy,* 94, 953-1001.

Kleidon, A. W., and J. L Koski, 1992, "Are Stock Prices Excessively Sensitive to Current Information?: Comment," *Journal of Economic Behaviour and Organization* forthcoming.

Kreps, D. M., and E. L. Porteus, 1978, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica,* 46,185-200.

I.eRoy, S. F., 1984, "Efficiency and the Variability of Asset Prices," *American Economic Review,* 74,183-187.

LeRoy, S. F.. 1989, "Efficient Capital Markets *and Martingales," Journal of Economic Literature,* 27, 1583-1621.

LeRoy, S. F., 1990, "Capital Market Efficiency: An Update," *Economic Review, Federal Reserve Bank of San Francisco.*

LeRoy, S.F., and C. J. LaCivita. 1981. “Risk Aversion and the Dispersion of Asset Prices.” *Journal of Business,* 54, 535-547.

LeRoy, S. P., and W. R. Parke, 1998, “Stock Price Volatility: Tests Based on the Geometric Random Walk,” working paper, University of California, Santa Barbara.

LeRoy, S. P., and R. D. Porter. 1981, “The Present-Value Relation: Tests Based on Implied Variance Bounds,” *Ecomometica,* 49, 555-574.

Mankiw, N. G., D. Romer, snd M. D. Shapiro, 1985, “An Unbiased Reexamination of Stock Market Volatility,” *Journal of Finance,* 40, 677-687.

Mankiw, N. G., D. Romer, and M. D. Shapiro, 1991, “Stock Market Forcastibility and Volatility: A Statistical *Appraisal,” Review of Economic Studies* 58, 455-477.

Marsh, T. A., and R. C. Merton, 1986, “Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices,” *American Economic Review,* 76, 483-498.

Mehra, R., and E. C. Prescott, 1985, “The Equity Premium: A Puzzle,” *Journal of Monerary Economics,* 15, 143-161.

Merton, R. C., 1987. “On the Current State of the Stock Market Rationality Hypothesis,” In S. Fischer *(cd.), Macroeconomics and Finance: Essays in Honor of Franco Modigliani,* MIT. Cambridge, Mass.

Michener, R. W., 1982, “Variance Bounds in a Simple Model of Asset Pricing,” *Journal of Political Economy,* 90, 166-175.

Poterba, J., and L. H. Summers, 1988, “Mean Reversion In Stock Prices: Evidence and Implications.” *Journal of Financial Economics* 22, 27-59.

Scott, I.., 1985, “The Present Value Model of Stock Market Prices Regression Tests and Monte Carlo Results,” *Review of Economics and Statistics,* 57, 599-605.

Shea, G. S., 1989, “Ex-Post Rational Price Approximations and the Empirical Reliability of the Present-Value Relation,” *Journal of Applied Economics* 4, 139-159.

Shiller, R. J.. 1979. “The Volatility of Long Term Interest Rates and Expectations Models of the Term Structure,” *Journal of Political Economy,* 87, 1190-1209.

Shlller, R. J., 1981, “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?” *American Economic Review,* 71, 421-436.

Shiller, R. J., 1984, “Stock Prices and Social Dynamics.” *Brookings Papers on Economic Activity,* 457-498.

Shiller, R. J., 1988, “The Probability of Gross Violations of a Present Value Variance Inequality,” *Journal of Political Economy,* 96, 1089-1092.

Singleton, K.J., 1988, “Expectations Models of the Term Structure and Implied Variance Bounds,” *Journal of Political Economy,* 88, 1159-1176.

Singleton, K. J., 1987, “Specification and Estimation of Intertemporal Asset Pricing Models,” In B. Friedman and F. Hahn (eds.). *Handbook of Monetary Economics,* North-Holland, Amsterdam.

Well, P., 1989, “The Equity Premium Puzzle and the Risk-Free Rate Puzzle,” *Journal of Monetary Economics,* 24, 401-421.

West, K. D.. 1987, “A Specification Test for Speculative Bubbles,” *Quarterly Journal of Economics,* 102, 553-388.

West, K. D., 1988a, “Dividend Innovations and Stock Price Volatility,” *Econometrica,* 56, 37-61.

West, K. D., 1988b, “Bubbles, Fads and Stock Price Volatility Tests: A Partial Evalutation,” *Journal* of Finance, 43, 636-656.