Some New Variance Bounds for Asset Prices

When equity prices are determined as the discounted sum of current and expected future dividends, Shiller (1981) and LeRoy and Porter (1981) derived a relationship between the variance of the price of equities, $p_t$, and the variance of the ex post realized discounted sum of current and future dividends: $p_t^*$; $\text{Var}(p_t^*) \geq \text{Var}(p_t)$. The literature has long since recognized that this variance bound is valid only when dividends follow a stationary process. Others, notably West (1988), derive variance bounds that apply when dividends are nonstationary. West shows that the variance in innovations in $p_t$ must be less than the variance of innovations in a forecast of the discounted sum of current and future dividends constructed by the econometrician, $\hat{p}_t$. Here we derive a new variance bound when dividends are stationary or have a unit root, that sheds light on the discussion in the 1980s of the Shiller variance bound: $\text{Var}(p_t - p_{t-1}) \geq \text{Var}(p_t^* - p_{t-1}^*)$! We also derive a variance bound related to the West bound: $\text{Var}(\hat{p}_t - \hat{p}_{t-1}) \geq \text{Var}(p_t - p_{t-1})$.

JEL codes: G12, G14

Keywords: volatility bounds, variance bounds, asset price variability.

Shiller (1981) and LeRoy and Porter (1981) proposed a test for “excess volatility” of stock prices, when these prices are determined as a discounted sum of current and expected future dividends: the variance of the equity price, $p_t$, should be less than the variance of the ex post realized discounted sum of dividends, $p_t^*$. Subsequently, Marsh and Merton (1986), Kleidon (1986), and Durlauf and Phillips (1988) criticized these tests, arguing that the test requires that the stochastic process for dividends be stationary. Here we demonstrate that if dividends are stationary or have a unit root, $\text{Var}(p_t - p_{t-1}) \geq \text{Var}(p_t^* - p_{t-1}^*)$. That is, expressing prices in first-differences, the Shiller–LeRoy–Porter inequality is reversed.

I thank Ken West, John Campbell, Jim Hodder, Allan Kleidon, and Stephen LeRoy for useful comments. This research was supported in part by a grant from the NSF to University of Wisconsin.

Charles Engel is from the Departments of Economics and Finance, University of Wisconsin, and NBER (E-mail: cengel@ssc.wisc.edu).

Received November 9, 2004; and accepted in revised form December 15, 2004.

Journal of Money, Credit, and Banking, Vol. 37, No. 5 (October 2005)
Copyright 2005 by The Ohio State University
In a sense, the profession long ago resolved how to implement variance bounds when dividends are nonstationary. Mankiw, Romer, and Shapiro (1985), and West (1988) introduce volatility bounds that are valid when there is a unit root in the dividend process. The West test involves a forecast of the discounted sum of current and future dividends constructed by the econometrician (a forecast based on a smaller information set than the market’s), \( \hat{\rho}_t \). Under the assumption that the econometrician has less information than markets, West shows that the variance in innovations in \( p_t \) should be less than the variance in innovations in \( \hat{\rho}_t \). Here we also derive a variance bound that is similar to that of West (1988): \( \text{Var}(\hat{\rho}_t - \hat{\rho}_{t-1}) \geq \text{Var}(p_t - p_{t-1}) \).

It has been noted (by Frankel and Stock, 1987 and Durlauf and Phillips, 1988) that the variance bound is a weaker restriction than imposed by the standard Euler equation. But it has been argued that variance bounds are nonetheless interesting because they provide some insight into why the Euler condition might fail. For example, Campbell, Lo, and MacKinlay (1997, p. 277) state, “The justification for using a variance-bounds test is not increased power; rather it is that a variance-bounds test helps one to describe the way in which the null hypothesis fails.”

In that spirit, there may be some value in re-examining the Shiller–LeRoy–Porter variance-bounds test. Shiller (1991) in particular argues for the intuitive appeal of his bound, \( \text{Var}(p^*_t) \geq \text{Var}(p_t) \), by asking readers to examine graphs of \( p^*_t \) and \( p_t \). As Shiller says (p. 421), “One is struck by the smoothness and stability of the ex post rational price series \( p^*_t \) when compared with the actual price series.” Flavin (1983), and especially Kleidon (1986), argued that this interpretation of the graphs was inconclusive. Just because \( p^*_t \) appears smoother does not mean it has lower variance, when the dividend process is very persistent. The subsequent exchange between Shiller (1988) and Kleidon (1988) demonstrates that the issue was not fully resolved. The result in this paper formalizes the observation that volatility of \( p_t \) compared to \( p^*_t \) does not imply that the present-value model is violated, in a very simple way. Given the near-random-walk behavior of stock prices, the “volatility” of the stock price is captured by \( \text{Var}(p_t - p_{t-1}) \); but the high volatility of the actual stock price is not inconsistent with the smooth behavior of \( p^*_t \) because we show here that the present-value model implies \( \text{Var}(p_t - p_{t-1}) \geq \text{Var}(p^*_t - p^*_{t-1}) \).

This observation should not, however, revive hope for the contention that stock prices are not excessively volatile. Both Mankiw, Romer, and Shapiro (1985), and West (1988) find that their variance bounds are violated in data for U.S. stock prices. We shall argue presently that the results of West (1988) should persuade us that the second variance bound derived here, \( \text{Var}(\hat{\rho}_t - \hat{\rho}_{t-1}) \geq \text{Var}(p_t - p_{t-1}) \), will also fail.

1. DEFINITIONS AND ASSUMPTIONS

\[ p^*_t \equiv \sum_{j=0}^{\infty} b^j d_{t+j}, \quad 0 < b < 1. \]  
\[ p^*_t \] is the “perfect foresight” price and \( d_t \) is the dividend at time \( t \).
\[ p_t \equiv E(p_t^* \mid I_t) \] is the information set of the market and \( p_t \) is the market price. 
\[ \hat{p}_t \equiv E(p_t^* \mid H_t) \] is an information set, \( H_t \subseteq I_t \).
\[ e_t \equiv p_t - E(p_t \mid I_{t-1}) \] and \( \hat{p}_t \equiv \hat{p}_t - E(\hat{p}_t \mid H_{t-1}) \).
\[ \sigma^2 \equiv \text{Var}(e_t), \, \hat{\sigma}^2 = \text{Var}(\hat{f}_t) \]

As in West (1988), we assume \( I_t \) is a linear space, spanned by the current and past values of a finite number of random variables, and \( I_t \subseteq I_{t+1} \). After \( s \) differences, all the random variables in \( I_t \) jointly follow a stationary ARMA(\( q,r \)) process for finite \( s, q, \) and \( r \).

Assume that at a minimum, \( H_t \) contains current and past values of \( d_t \).
West (1988) shows \( p_t^* - p_t = \sum_{j=1}^{\infty} b^j e_{t+j} \) and \( p_t^* - \hat{p}_t = \sum_{j=1}^{\infty} b^j f_{t+j} \).

**COMMENT:** Shiller (1981) shows \( \text{Var}(p_t^*) \geq \text{Var}(p_t) \) when \( d_t \) is stationary.
West (1988) shows \( \hat{\sigma}^2 \geq \sigma^2 \) when \( d_t \) is a linear process integrated of any order.

**PROPOSITION 1:** Suppose \( d_t \) is \( I(1) \) or \( I(0) \), and all of the assumptions above hold. Then

\[ \text{Var}(p_t - p_{t-1}) \geq \text{Var}(p_t^* - p_{t-1}^*) . \]

**PROOF:**
\[ \text{Var}(p_t - p_{t-1}) = \text{Var}(\{ E(p_t \mid I_{t-1} - p_{t-1} \} + \{ p_t - E(p_t \mid I_{t-1}) \} = \Gamma + \sigma^2 , \]
where \( \Gamma \equiv \text{Var}(E(p_t \mid I_{t-1} - p_{t-1}) \). The last equality holds because \( e_t \equiv p_t - E(p_t \mid I_{t-1}) \) is uncorrelated with \( t - 1 \) information.

\[ \text{Var}(p_t^* - p_{t-1}^*) = \text{Var}(\{ E(p_t^* \mid I_{t-1} - p_{t-1} \} + \{ p_t - E(p_t^* \mid I_{t-1}) \} + \sum_{j=1}^{\infty} b^j e_{t+j} - \sum_{j=1}^{\infty} b^j e_{t+j-1} = \text{Var}(\{ E(p_t^* \mid I_{t-1} - p_{t-1} \} + (1 - b)e_t + (1 - b)\sum_{j=1}^{\infty} b^j e_{t+j} = \Gamma + (1 - b)^2 \sigma^2 + \frac{(1 - b)^2 b^2}{1 - b^2} \sigma^2 = \Gamma + \frac{1 - b}{1 + b} \sigma^2 \leq \text{Var}(p_t - p_{t-1}) . \]

**COMMENT:** Note the surprising relationship to the Shiller (1981) variance bound. Also note that the proposition does not extend to the claim for all \( k > 0 \), \( \text{Var}(p_{t+k} - p_t) \geq \text{Var}(p_{t+k}^* - p_t^*) \). (See the Appendix in the working paper version of this paper for counterexamples for \( k > 1 \).)

**DISCUSSION OF PROPOSITION 1:** For convenience, define \( \Delta p_t \equiv p_t - p_{t-1} \) and \( \Delta p_t^* \equiv p_t^* - p_{t-1}^* \). We see from the definitions above that \( E(\Delta p_t \mid I_{t-1}) = E(\Delta p_t^* \mid I_{t-1}) \). That is, \( E(\Delta p_t \mid I_{t-1}) \) is an unbiased (relative to the information set \( I_{t-1} \)) forecast of both \( \Delta p_t \) and \( \Delta p_t^* \). So, we can write:

\[ \text{Var}(p_t^* - p_{t-1}^*) = \text{Var}(E(\Delta p_t \mid I_{t-1})) + \text{Var}(p_t^* - p_{t-1}^* - E(\Delta p_t \mid I_{t-1})) , \]
and
\[ \text{Var}(p_t - p_{t-1}) = \text{Var}(E(\Delta p_t \mid I_{t-1})) + \text{Var}(p_t - p_{t-1} - E(\Delta p_t \mid I_{t-1})) . \]

Proposition 1, which states \( \text{Var}(p_t - p_{t-1}) \geq \text{Var}(p_t^* - p_{t-1}^*) \), is equivalent then to the statement, \( \text{Var}(p_t - p_{t-1} - E(\Delta p_t \mid I_{t-1})) \geq \text{Var}(p_t^* - p_{t-1}^* - E(\Delta p_t \mid I_{t-1})) . \)

That is, the market at time \( t - 1 \), which has information \( I_{t-1} \), can make a better forecast of \( \Delta p_t^* \) than of \( \Delta p_t \) (Here, “better forecast” means a forecast error with lower variance.)

To understand this, first we see that of course the forecast error the market makes for \( \Delta p_t \) is just its forecast error for \( p_t \), since \( p_t \) is in \( I_{t-1} \). That is,
\[ \text{Var}(p_t - p_{t-1} - E(\Delta p_t \mid I_{t-1})) = \text{Var}(p_t - E(p_t \mid I_{t-1})) = \text{Var}(e_t) = \sigma^2 . \]

But in forecasting \( \Delta p_t^* \), we must recognize that neither \( p_t^* \) nor \( p_{t-1}^* \) are in \( I_{t-1} \). The forecast errors for \( p_t^* \) and \( p_{t-1}^* \) are correlated—indeed they are perfectly correlated (as we show shortly). So, while the variance of the market’s forecast error of \( p_t^* \) is greater than the variance of the market’s forecast error of \( p_t \), the variance of the market’s forecast error of \( p_t^* - p_{t-1}^* \) is much smaller than the variance of its forecast error of \( p_t^* \)—and, as Proposition 1 implies, even smaller than the variance of the forecast error of \( \Delta p_t \).

To see this, use the fact that, from the definitions above, \( p_t^* - p_{t-1} \) and \( p_t - p_{t-1} \) satisfy the following relationships:
\[ p_t^* - p_{t-1} = d_{t-1} + bp^*_t, \]
\[ p_t - p_{t-1} = d_{t-1} + bE(p_t \mid I_{t-1}). \]

Subtraction gives us \( p_t^* - p_{t-1} = b(p_t^* - E(p_t \mid I_{t-1})). \) \( p_t^* - p_{t-1} \) is the market’s forecast error of \( p_t^* \) at time \( t - 1 \), and \( p_t^* - E(p_t \mid I_{t-1}) \) is the market’s forecast error of \( p_t^* \) at time \( t - 1 \). This shows that the forecast errors of \( p_t^* - p_{t-1} \) and \( p_t - p_{t-1} \) based on \( I_{t-1} \) are perfectly correlated.

The variance of the market’s forecast error of \( p_t^* \) is given by
\[ \text{Var}(p_t^* - E(p_t \mid I_{t-1})) = \frac{1}{b^2} \text{Var}(p_t^* - p_{t-1}) = \frac{1}{b^2} \sum_{j=1}^{\infty} b^j e_{t+j-1} = \frac{1}{1 - b^2} \sigma^2 . \]

Clearly the variance of the market’s forecast error of \( p_t^* \) is greater than the variance of the market’s forecast error of \( p_t \). But, now consider the variance of the forecast error of \( \Delta p_t^* \):
\[ \text{Var}(p_t^* - p_{t-1}^* - E(\Delta p_t \mid I_{t-1})) = \text{Var}([p_t^* - E(p_t \mid I_{t-1})] - [p_{t-1}^* - p_{t-1}]) \]
\[ = \text{Var}((1 - b)[p_t^* - E(p_t \mid I_{t-1})]) = \frac{(1 - b)^2}{1 - b^2} \sigma^2 \]
\[ = \frac{1 - b}{1 + b} \sigma^2 . \]

The variance of the forecast error of \( \Delta p_t^* \) is less than the variance of the forecast error of \( p_t^* \) and \( p_t \).
The intuition of Proposition 1 discussed here is in many ways similar to Kleidon’s (1986) discussion of why it is misleading to draw inferences from the fact that the graph of \( p_t^* \) in Shiller (1981) is smoother than the graph of \( p_t \). However, Kleidon did not consider models in which dividends could follow general \( I(1) \) processes and did not examine the variances of differences in prices, so the analogy to that discussion is imperfect.

**PROPOSITION 2:** Suppose \( d_t \) is \( I(1) \) or \( I(0) \), and all of the above assumptions hold. Then

\[
\text{Var}(\hat{p}_t - \hat{p}_{t-1}) \geq \text{Var}(p_t - p_{t-1}) .
\]

**PROOF:** Following the same steps as above, but replacing \( p_t \) with \( \hat{p}_t \), we have \( \text{Var}(\hat{p}_t - \hat{p}_{t-1}) \) = \( \hat{\Gamma} + \hat{\sigma}^2 \), where \( \hat{\Gamma} \equiv \text{Var}(E(p_t \mid H_{t-1} - p_{t-1}) \) and \( \text{Var}(p_t^* - p_{t-1}^*) = \hat{\Gamma} + ((1 - b)/(1 + b))\hat{\sigma}^2 \). It follows that

\[
\hat{\Gamma} + \frac{1 - b}{1 + b} \hat{\sigma}^2 = \Gamma + \frac{1 - b}{1 + b} \sigma^2 .
\]

Then,

\[
\text{Var}(p_t - p_{t-1}) = \Gamma + \sigma^2 = \Gamma + \frac{1 - b}{1 + b} \sigma^2 + \frac{2b}{1 + b} \sigma^2 \leq \hat{\Gamma} + \frac{1 - b}{1 + b} \hat{\sigma}^2 + \frac{2b}{1 + b} \sigma^2 = \text{Var}(\hat{p}_t - \hat{p}_{t-1}) .
\]

The inequality in this expression follows because West (1988) shows that \( \hat{\sigma}^2 \geq \sigma^2 \).

**COMMENT:** Note the relationship of this variance bound to that of West (1988). At first glance, one might think that the two propositions contain the same result in the special case in which \( d_t = d_{t-1} + e_t \). That is true, but only trivially. Because both Proposition 1 of West (1988) and Proposition 2 here assume \( H_t \) includes current and past values of \( d_t \), we have in this case that \( \hat{\sigma}^2 = \sigma^2 \), and \( \text{Var}(\hat{p}_t - \hat{p}_{t-1}) = \text{Var}(p_t - p_{t-1}) \). That is, any information in \( I_t \) that is not in \( H_t \) is not helpful in forecasting \( d_{t+1} \).

**DISCUSSION OF PROPOSITION 2:** Think of \( \hat{p}_t \) as the forecast an econometrician makes of \( p_t^* = \sum_{j=0}^{\infty} b^j d_{t+j} \), based on a VAR as in West (1988).

Consider the relationship between the forecast of \( \Delta \hat{p}_t \equiv \hat{p}_t - \hat{p}_{t-1} \) and \( \Delta p_t \). Following the same logic as in the Discussion of Proposition 1, we can write

\[
\text{Var}(p_t - p_{t-1}) = \text{Var}(E(\Delta p_t \mid H_{t-1})) + \text{Var}(p_t - p_{t-1} - E(\Delta p_t \mid H_{t-1})) ,
\]

\[
\text{Var}(\hat{p}_t - \hat{p}_{t-1}) = \text{Var}(E(\Delta \hat{p}_t \mid H_{t-1})) + \text{Var}(\hat{p}_t - \hat{p}_{t-1} - E(\Delta \hat{p}_t \mid H_{t-1})) ,
\]

where we have used the fact that \( E(\Delta p_t \mid H_{t-1}) = E(\Delta \hat{p}_t \mid H_{t-1}) = E(\Delta p_t^* \mid H_{t-1}) \). The theorem then implies that
Var\left(p_t - p_{t-1} - E(\Delta p_t \mid H_{t-1})\right) \leq \text{Var}(\hat{p}_t - \hat{p}_{t-1} - E(\Delta \hat{p}_t \mid H_{t-1})).

Notice the comparison to the West’s (1988) result. Since \( p_{t-1} \) is in \( I_{t-1} \) and \( \hat{p}_{t-1} \) is in \( H_{t-1} \), we can write West’s result as:

\[ \text{Var}\left(p_t - p_{t-1} - E(\Delta p_t \mid I_{t-1})\right) \leq \text{Var}(\hat{p}_t - \hat{p}_{t-1} - E(\Delta \hat{p}_t \mid H_{t-1})) \]

Another related paper is that of Engel and West (2004). They show that as \( b \to 1 \), \( \text{Var}[(1 - b)(\hat{p}_t - \hat{p}_{t-1})] = \text{Var}((1 - b)(p_t - p_{t-1})) \). Their proof, however, takes a very different tack than the proofs here. They show that as \( b \to 1 \), \( \text{Var}[(1 - b)(\hat{p}_t - E(\hat{p}_t \mid H_{t-1}))] = \text{Var}((1 - b)E(p_t \mid I_{t-1})) \). They then use the result from Engel and West (2005) that as \( b \to 1 \), \( \hat{p}_t - E(\hat{p}_t \mid H_{t-1}) \approx \hat{p}_t - \hat{p}_{t-1} \) and \( p_t - E(p_t \mid I_{t-1}) = p_t - p_{t-1} \) to conclude that \( \text{Var}((1 - b)(\hat{p}_t - \hat{p}_{t-1})) = \text{Var}((1 - b)(p_t - p_{t-1})) \) when \( b \) is near one.

Now consider the relationship between the variance of \( \Delta \hat{p}_t \) and \( \Delta p_t^s \). Proposition 2, combined with Proposition 1, gives us

\[ \text{Var}(\hat{p}_t - \hat{p}_{t-1}) \geq \text{Var}(p_t - p_{t-1}) \geq \text{Var}(p_t^s - p_{t-1}^s). \]

This means that the variance of \( \Delta \hat{p}_t \) is an upper bound on the variance of \( \Delta p_t^s \). Even if the present-value model is not how the market prices equities, the econometrician can still calculate an upper bound on the variance of the change in the ex post discounted sum of current and future dividends.

As we have noted, the graphs of Shiller (1981) in essence confirm that the variance bound of Proposition 1 is satisfied. However, the results of West (1988) in essence confirm that the variance bound of Proposition 2 is not satisfied. The near-random-walk behavior of equity prices means that \( \text{Var}(p_t - E(p_t \mid I_{t-1})) \) will not be too different than \( \text{Var}(p_t - p_{t-1}) \). Also, West’s estimates show that dividends are nearly a random walk, suggesting that \( \text{Var}(\hat{p}_t - E(\hat{p}_t \mid H_{t-1})) \) is none too different than \( \text{Var}(\hat{p}_t - \hat{p}_{t-1}) \). Given the gross violations of the bound \( \text{Var}(\hat{p}_t - E(\hat{p}_t \mid H_{t-1})) > \text{Var}(p_t - E(p_t \mid I_{t-1})) \) that West reports, we can quite confidently hazard the guess that the bound \( \text{Var}(\hat{p}_t - \hat{p}_{t-1}) \geq \text{Var}(p_t - p_{t-1}) \) will also fail.

LITERATURE CITED


