ECONOMETRIC REVIEWS, 5(2), 273-278 (1986)

COMMENT

This paper (hereafter "MM") attempts to establish some finite sample properties of some time series estimators of a simple present value model, mostly in nonstationary environments. The subject is important and topical. The paper obviously required an enormous amount of effort. Regrettably, there is no evidence that the results have any practical relevance.

The most important problem, which potentially affects every statistic that MM report, is that the choice of parameters for MM's artificial data is neither defended nor explained. No evidence is presented that their artificial data is qualitatively similar enough to any real world data to make MM's results of interest to the applied researcher. In fact, such evidence as I have been able to establish suggests precisely the opposite.

I will elaborate on this mostly in connection with MM's simple integrated model (experiment 5). My basic point here is that MM have generated data that follow a random walk with essentially zero drift. Since it is well known that the resulting small sample distribution of their estimators is highly nonnormal, MM's result that there are "important small-sample biases in. . .test statistics" (section I) that assume an asymptotic normal distribution is, I'm afraid, neither very surprising nor very informative.

Before discussing this point in detail, let me emphasize that my aim is not to argue that the statistics have zero small sample bias. That would be surprising for many reasons, including that my own Monte Carlo work in West (1986, 1986b) has suggested some bias. My point rather is that MM have done little to help quantify the

0747-4938/86/0502-0273$3.50/0

bias, and certainly their section V use of the Monte Carlo results as guidelines for the exact small sample distribution of the tests applied to real world data is unsupported.

MM state in their appendix that many parameters "were selected through reference to. . .the S and P 500 (1889-1979)." Unfortunately, this statement manifestly does not apply to their artificial X_t data in their simple integrated model. The parameters C_x, ϕ_1, ϕ_2, and ϕ_3 reported in experiment 5 in appendix 1 do in fact match the estimates I reported in my 1985 and 1986b papers (incidentally, for 1874-1980, not 1889-1979). But given $\delta_0 = \sqrt{10}$ and $\sigma(\varepsilon_x) = 1$, the implied process for the artificial X_t series is quite different from that estimated for the actual S and P dividend series. Consider in particular the ratio of the mean to the standard deviation (μ/sd) of ΔX_t. This is .034/3.6 $\cong$ .01 for their artificial data. (If $\phi_1 = 1$ and $\lambda = \phi_2 = \phi_3 = 0$, $\text{var}(\Delta X_t) = \delta_0^2 \text{var}(z_t) + \text{var}(\varepsilon_{xt}) = 10.1 + 1 = 11$; the values of λ and the ϕ_i actually used imply a little positive serial correlation in ΔX_t, and result in a variance of about 12.98.) The estimated value for the S and P, however, is .034/.40 $\cong$ .09, a figure that is higher by a factor of nine. The point estimate of μ/sd for the S and P price data is also higher than that of MM's artificial data, by a factor of about four to five.

Now, we know that the small sample distribution of MM's estimators will be nonnormal if $\mu = 0$, and, at least in estimation of a unit autoregressive root, will be normal as $\mu/\text{sd} \to \infty$ (Evans and Savin (1984)). The question, then, is how well MM's evidence of bias for μ/sd $\cong$ .01 applies when μ/sd $\cong$ .09. That the MM results may have little relevance is suggested by the following.

I repeated the Monte Carlo simulation in West (1986b) of the point estimate (but not, for computational simplicity, the standard error) of MM's equation (33). (It is this simulation that, I believe, MM are referring to in their footnote 14, since there are no simulations in the 1984b version of my paper.) The simulations assumed a simple present value model with X_t following a pure random walk with drift. Under the null hypothesis, the population value of equation (33) is zero; in small samples, an ideal test procedure will produce an equal number of positive and negative figures. Whether the procedure has a spurious tendency to produce negative figures is important, since under certain circumstances negative figures that are large in absolute value can be interpreted as evidence against the simple present value model (West (1986b)).

As in MM, I used b = .9615, and threw out any random sample that produced $b < .7$ or $b > 1.0$. Because of computer constraints, only 200 samples, T = 100 (plus initial lag and terminal lead), were drawn. σ_v^2 was obtained from the residuals from a regression of ΔX_t on a constant. (N.B.: In contrast to MM, in this Monte Carlo study as in West (1986b) I imposed unit roots when these were assumed present.) See West (1986b) for further details on the procedure.

First, I followed my previous study and set μ/sd for ΔX_t to match the estimates from the S and P. As reported in line A of Table I, 52 percent of the resulting samples yielded a negative point estimate for (33). This is slightly more than the ideal 50 percent figure. Then I followed MM and set μ/sd to match their artificial data in experiment 5. Using the same random number seed, the percentage of negative estimates jumped to 78 (line B). (Line C is explained below. Incidentally, the mean value of (33) was positive for both line A and line B, apparently because of a few large positive outliers.)

This simple experiment indicates that, at least for their simple integrated model, MM's small sample results most likely reflect nothing other than what is probably a wildly unrealistic parameterization of the artificial data. This is suggested as well by Dickey and Fuller (1981). Dickey and Fuller consider testing for a unit root using an equation like (25) (i.e., one without a trend term as a regressor). Table VII in Dickey and Fuller (1981) indicates that with a sample size of 100, a nominal .05 Dickey-Fuller test for a unit root has an actual size of .26 when μ/sd = .5; the INT(X) entry in column 5 of MM's Table 2 apparently indicates an actual size of .028 for μ/sd $\cong$ .01. It is not clear that

TABLE I

Estimation of (33) in Monte Carlo Experiment

	Percentage of estimates that are negative
A. Mean/s.d. of ΔX_t match S and P estimates, unit AR root imposed	52
B. Mean/s.d. of ΔX_t match Mattey and Meese's experiment 5, unit AR root imposed	78
C. Mean/s.d. of ΔX_t match Mattey and Meese's experiment 5, unit AR root not imposed	96

Note: Experiment is based on 200 replications, with 11-13 aborts. Lines A and B use the procedure applied to the actual S and P data in West (1986b), line C uses the procedure applied by Mattey and Meese. An ideal procedure produces a figure of 50.

Dickey and Fuller's evidence is less relevant to studies using the S and P data than is MM's. In fact, Dickey and Fuller's evidence is almost certainly more relevant to most studies using macroeconomic data since these data typically are much less noisy than are stock data. In MM's consumption data, for example, the estimated μ/sd is about .36 for $\Delta(\log C_t)$, about .47 for ΔC_t.

More generally, for their other experiments as well, MM present no evidence that their artificial data is close enough to real world data to make the experiments of interest. Unless I accidentally missed the relevant passages, MM do not even tell us, for example, the medians, means, standard deviations, auto- or cross-correlations of levels or first differences of price, X_t or returns, for any of their experiments. The only bit of information that I noticed relates to the bubbles experiment. We are told in the concluding section that this experiment on average produces a certain ratio that is over a thousand times larger than that estimated for the S and P. This is not the kind of remark that gives one confidence that the artificial data is qualitatively similar to real world data.

Before concluding, I have space to discuss only one more aspect of MM's study. This is their puzzling decision not to impose known unit roots in estimation, perhaps a consequence of a too literal reading of West (1986). The problem with this of course is that there is a downward bias in the point estimates of an AR coefficient if there is a unit root (Evans and Savin, 1984). Some indication of the resulting loss of efficiency is perhaps indicated in Line C of Table I. I repeated the Monte Carlo simulation reported in line B and described above. This time, instead of regressing ΔX_t on a constant, I followed MM's WEST1 and WEST2 procedures and regressed X_t on a constant and X_{t-1}. Together with the estimate of β, the resulting estimate of the coefficient on X_{t-1} was used to calculate the equation (33) parameter that MM call α_1. As indicated in line C, the percentage of negative estimates jumped to 96. This is reasonably close to the 90-95 percent produced by MM's more complicated experiment (MM, Table 10). (The WEST1 entry in MM's Table 10 is negative for Q90, positive for Q95. Since the sign of the asymptotically N(0,1) test statistic is the same as that of the point estimate of (33), some 90-95 percent of MM's estimates of (33) were negative. I go through this explanation mainly to point out that MM's footnote 14 is misleading.)

To sum up Table I: The percentage of negative estimates produced by the procedure used in West (1986b) appears to be quite near its ideal value of 50, under the null hypothesis of a simple present value model, and for a dividend series whose μ/sd is matched to that estimated for the actual S and P series (line A). Under the same null hypothesis, the percentage produced by the procedure that MM call WEST1 is much greater than 50, for dividend data whose μ/sd matches that of MM's simple integrated model (line C). I conclude from this detailed examination of how MM generated the statistic for WEST1, in their simple integrated model, not only that MM have yet to argue that their results are important for the interpretation of empirical studies they cite, but that the evidence to date suggests that their results have little practical relevance. This conclusion is reinforced by the Monte Carlo simulations of the simple present value in West (1986). These indicated that for sample sizes of 100, with X_t following a random walk with drift, the asymptotic normal approximation can be quite good, even for data as noisy as stock prices.

Mattey and Meese have presented us with what must have been a backbreaking effort, in a study of an important topic. I am sorry that their claim to have quantified the power and small sample bias of some recently developed test procedures is unsupported.

Kenneth D. West
Princeton University

ACKNOWLEDGEMENTS

I thank the National Science Foundation for financial support. This comment was first drafted while I was a National Fellow at the Hoover Institution.

ADDITIONAL REFERENCES

West, K.D. (1985). "A Specification Test for Speculative Bubbles," Princeton University Woodrow Wilson School Discussion Paper No. 97, July 1985.

West, K.D. (1986b). "Dividend Innovations and Stock Price Volatility," National Bureau of Economic Research Working Paper No. 1833, March 1986.