NBER TECHNICAL WORKING PAPER SERIES

# THE DIVIDEND RATIO MODEL AND SMALL SAMPLE BIAS: A MONTE CARLO STUDY

John Y. Campbell

Robert J. Shiller

Technical Working Paper No. 67

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1988

The authors thank Gary Shea for helpful discussions. This research was supported by the National Science Foundation. This reasearch is part of NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER TECHNICAL WORKING PAPER #67
July 1988

# THE DIVIDEND RATIO MODEL AND SMALL SAMPLE BIAS: A MONTE CARLO STUDY

## ABSTRACT

Small sample properties of parameter estimates and test statistics in the vector autoregressive dividend ratio model (Campbell and Shiller [1988 a,b]) are derived by stochastic simulation. The data generating processes are cointegrated vector autoregressive models, estimated subject to restrictions implied by the dividend ratio model, or altered to show a unit root.

John Y. Campbell
Woodow Wilson School
Princeton University
Princeton, NJ

Robert J. Shiller
Cowles Foundation
Yale University
Box 2125 Yale Station
New Haven, CT 06520

In this paper, we evaluate the estimates and tests of the present value model in the log linear vector autoregressive dividend ratio model formulation of Campbell and Shiller [1988a,b]. The data generating processes for the simulations are cointegrated vector autoregressive processes that are consistent with the efficient markets model. An estimated vector autoregressive model and an alternative model with an imposed unit eigenvalue are both used.

## I. The Dividend Ratio Model

The dividend ratio model (Campbell and Shiller [1988a,b]) states that the log dividend-price ratio $\delta_t$ (equal to $d_{t-1} - p_t$ where $d_{t-1}$ is the log dividends per share paid the period before time t and $p_t$ is the log price per share at the beginning of time t) is given by:

$$\delta_t = E_t \delta_t^* \tag{1}$$

$$\delta_t^* = \sum_{j=0}^{\infty} \rho^j [r_{t+j} - \Delta d_{t+j}] - k/(1-\rho) \tag{2}$$

where $r_t$ is the one-period discount rate at time t. The model was derived by linearizing the exact expression for $\delta_t$, and the constant k equals $\ln(1+\exp(\delta)) - \delta\exp(\delta)/(1+\exp(\delta))$ where $\delta$ is the point of linearization.

## II. The Data Generating Processes

Two data generating processes were estimated. The first was estimated subject to the constraint that the present value model holds with a constant discount rate. The second was estimated subject to the constraint that the present value model holds with discount rate varying through time with the commercial paper rate.

Let us consider the first data generating process. The vector for the vector autoregression was $z_t = [\delta_{t+1}, \Delta d_t, \epsilon 30_{t+1}]$ where $\epsilon 30_t$ is a log earnings price ratio based on a thirty year moving average of earnings, as in campbell and Shiller [1988b]. Let us suppose that $z_t$ is a Gaussian vector first-order autoregressive (AR1) process:

(3) $$z_t = Az_{t-1} + v_t.$$

As was shown in Campbell and Shiller [1988a], the efficient markets model (1) requires that:

(4) $$e1'(I - \rho A) = -e2'A$$

where $e1' = [1, 0, 0]$ and $e2' = [0, 1, 0]$. Equation (4) may be interpreted as requiring that the linearized one-period return $\xi_t = \delta_t - \rho\delta_{t+1} + \Delta d_t$ be unforecastable. The restriction (4) can be written in another way:

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

It is a feature of the Wald test that the test statistic depends on the way the restrictions are written. In our previous paper we referred to a Wald test of the restriction (4') as a test that $\delta_t = \delta'_t$ where $\delta'_t = -e2'A(I-\rho A)^{-1}z_t$ or as a test that infinite-period returns are unforecastable. Here, we refer to a Wald test using (4) as the linear Wald test of the model, since (4) is linear in the parameter matrix A, and to a Wald test using (4') as a nonlinear Wald test of the model.

An estimate of the vector autoregression parameters was derived by ordinary least squares subject to the restriction (4).[2] The estimated A and Σ using the full data set in Campbell and Shiller [1988b] (regressions containing 86 observations, with $\xi_t$ from 1901 to 1986) were:

$$A = \begin{bmatrix} 0.6101 & -0.2103 & 0.0862 \\ -0.4289 & -0.1969 & 0.0806 \\ 0.0079 & 0.1041 & 0.8741 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 0.0415 & 0.0076 & 0.0338 \\ 0.0076 & 0.0134 & -0.0025 \\ 0.0338 & -0.0025 & 0.0362 \end{bmatrix}$$

The (real) eigenvalues of A are 0.8681, 0.7254 and -0.3062. The largest is fairly close to one, so that its half life is five years. Such eigenvalues are to be expected, as the variables $\delta_t$ and $\epsilon 30_t$ show some persistence through time.[3]

The matrix Σ is not terribly ill-conditioned, its condition number (the ratio of largest to smallest eigenvalue) is 76.09. Thus, we are not in a

---

[2]By a change of variables, we define $\tilde{z}_t = Sz_t$ where S is the matrix:

$$S = \begin{matrix} 1 & 0 & 0 \\ \rho & -1 & 0 \\ 0 & -0 & 1 \end{matrix}$$

Now, $\tilde{z}_t = \tilde{A}\tilde{z}_t + \tilde{v}_t$ where $\tilde{A} = SAS^{-1}$. One imposes the restriction that $\xi_t$ is unforecastable on the coefficients of the second equation, estimates $\tilde{A}$ and $\tilde{\Sigma}$ and then recovers A and Σ.

[3]If we were using similar methods to evaluate a VAR-p model with p greater than 1, using the first order autoregressive companion form $z_t = Az_{t-1} + v_t$, then the restrictions (2) would imply that there is a zero eigenvalue for A. This is so because one of the rows of A would be el'. However, in the first order case the matrix A need not be singular.

situation in which there is a linear dependence among the rows of $z_t$, as would happen in the case of "no superior information." The no superior information case occurs when the log real price carries only information which is in the other variables in the regression. In that case the only information relevant to forecasting the future comes from dividend changes ($\Delta d_{t-1}$), the spread between the current log real dividend and the 30-year moving average of log real earnings ($\delta_t - \epsilon 30_t$), and their lagged values, so that log real price is linear in these and its innovations therefore linear in their innovations. The "no-superior information" case has figured prominently in monte carlo evaluations of volatility tests (see for example Kleidon [1986], Mattey and Meese [1986], Fama and French [1988]). In the present case, a component of innovations in $\delta_t$ uncorrelated with the other innovations feeds into future movements in dividends. (The moving average representation of the system would give weight to this component in the determination of future dividend changes). The example here can therefore be interpreted as one in which economic agents have some information relevant to predicting future dividends beyond dividends and earnings.

The second data generating process, which assumes that the present value model (1) holds where discount rates move through time with the commercial paper rate, was estimated in the same way, except that the vector $z_t = [\delta_{t+1}, \Delta d_t - r_t, \epsilon 30_{t+1}]$ (all variables demeaned) where $r_t$ is the prime commercial paper rate. In this vector, $\Delta d_t$ is measured in nominal terms; in effect the interest rate is used to deflate the change in dividends. With this change in data, the same constraints (4) or (4′) apply here. The least squares estimates of A and Σ were:

$$A = \begin{bmatrix} 0.5840 & -0.2735 & 0.1068 \\ -0.4539 & -0.2560 & 0.1000 \\ 0.0087 & 0.0824 & 0.8743 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 0.0410 & 0.0072 & 0.0338 \\ 0.0072 & 0.0129 & -0.0025 \\ 0.0338 & -0.0025 & 0.0363 \end{bmatrix}$$

which are fairly similar to those of the constant discount rate case.

## III. The Present Value Model Tests

The methods used in Campbell and Shiller [1988a,b] are to estimate the vector autoregression for $z_t$ and to test the restrictions (4) or (4') by an ordinary Wald test. The theoretical log ratio $\delta'_t = -e2'A(I-\rho A)^{-1}z_t$ is computed and compared with the actual dividend-price log ratio $\delta_t$. Finally, the theoretical linearized return $\xi'_t = \rho\delta'_{t+1} + \Delta d_t - \delta'_t$ is computed and compared with the actual linearized return $\xi_t$.

The Campbell-Shiller results [1988b] for the constant discount rate case are reproduced in Table 1, Panel A. The linear Wald test rejects the restrictions (4) at the 4.1% level, the nonlinear Wald test rejects at a much higher significance level (with a Wald statistic of 104.424 and three degrees of freedom, the computed significance level is $1.545\times10^{-9}$). The equation determining $\delta'_t$ in terms of the vector $z_t$ does not put a coefficient of 1.000 on $\delta_t$ and zero on the other two variables, as the efficient markets model requires. The equation determining $\delta'_t$ may be interpreted as an equation determining $p'_t$ (equal to $d_{t-1} - \delta'_t$): $p'_t = .776e30_t + 0.256p_t + 0.046d_{t-1} - 0.078d_{t-2}$ where $e30_t$ is a thirty-year moving average of real earnings. This equation does not put all of the weight in determining $p'_t$ on $p_t$ as efficient markets would require; rather it puts roughly three-quarters of the weight on $e30_t$ and only one quarter of the weight on the price $p_t$. We see that $\delta'_t$ is somewhat less variable than $\delta_t$ and has little correlation with it. The variable $\xi'_t$ is much less variable than $\xi_t$ but is highly correlated with it. These results were interpreted as indicating a substantial failure of the efficient markets model (1), in the direction of excess volatility for stock returns.

The stochastic simulation results, Table 1 Panel B, are generally supportive of this interpretation of the results. While the size of the 5% linear Wald test is really 8.6%, the rejection of the hypothesis with the actual data occurred at a significance level of 4.1%, and such a rejection occurred in 7.2% of the iterations. Thus, we can say in light of the stochastic simulations that the linear Wald test with our data rejects at roughly the 7% level. The size of the nonlinear Wald test is more problematic: the 5% nonlinear Wald test rejected 23.4% of the time. But the rejection of the nonlinear Wald test with the actual data was so dramatic, that in none of the 1000 iterations was the significance level of the Panel A results achieved. While there is a bias that puts some weight on the extraneous variable $e30_t$ in determining $p'_t$, the average weight put on $e30_t$ in the stochastic simulations is only 0.215, not, as in the estimated equation, 0.776. There is also a downward bias in the standard deviation of $\xi'_t$: the procedure tends to conclude that returns ought to be less variable than they are even when the model is correct. Both of these biasses might be interpreted as manifestations of the same general phenomenon reported by Flavin [1983], Kleidon [1986], Marsh and Merton [1986] and others. However, the extent of the bias is not enough to give a likely reconciliation between the model and the estimated regression. Not once in the 1000 iterations was an estimated standard deviation of $\xi'_t$ less than 0.3 times the standard deviation of actual $\xi_t$.

Another aspect of Table 1 is well worth noting. The stochastic simulations uniformly put a high correlation between $\delta_t$ and $\delta'_t$: the average correlation was 0.946 with a standard deviation of only 0.063. The actual correlation, at 0.175, was dramatically below this.

Table 2 shows the same results for the time-varying discount rate case. The results are very similar to those shown in Table 1. The rejection of the efficient markets model with the actual data was somewhat less dramatic than in the constant discount rate case: with the actual data the standard deviation of $\xi'_t$ was almost half that of $\xi_t$. In only 1.7% of the simulation iterations was the standard deviation of $\xi'_t$ less than half that of $\xi_t$.

While these results are quite favorable to the interpretation of results in Campbell and Shiller [1988b], it should be noted that alternative data generating processes are available that give substantially worse small sample performance to the estimators and tests. To show this, we altered the A matrix so that the largest eigenvalue equaled 1.000. This was done by increasing the single element A(3,3), a parameter that is not involved in the restrictions (4). To achieve a unit eigenvalue, the parameter A(3,3) had to be increased by 1.413 standard errors in the constant discount rate case, not an implausible amount. Note that the small sample properties of the tests diverge even more from those suggested by our asymptotic distribution theory that assumed stationarity for $z_t$. Most strikingly, the nonlinear Wald test rejects more than half the time at the 5% level. Still, rejections of the efficient markets model in the constant discount rate case at the significance level found in the actual data occurred less than one percent of the time. In none of the 1000 iterations was the the estimated standard deviation of $\xi_t'$ less than .3 times the actual standard deviation, as observed with the actual data in Table 1 Panel A.

With the time-varying discount rate case (table 4), the unit eigenvalue process small sample properties show about as much discrepancy from the asymptotic properties as we saw in Table 3. In some respects, the situation looks somewhat worse: 13.4% of the time the nonlinear Wald test rejects at the significance level we observed with the actual data, and 23.8% of the time the estimated standard deviation of $\xi_t'$ is less than half that of $\xi_t$. Here, we do not have a very significant rejection of the efficient markets model.

The unit eigenvalue case that was the basis of Tables 3 and 4 is an extreme one. It implies that dividend-price ratios, earnings-price ratios, and dividend growth rates are all nonstationary stochastic processes. It has not been suggested in the literature that these are nonstationary. We believe that the encouraging results in tables 1 and 2 are more likely to be relevant to the actual data, which appear to be stationary.

Table 1

Vector Autoregression Results - Constant Discount Rate (Real Returns)

A. Actual Results (from Campbell and Shiller Table 2 Panel A)

Linear Wald Test (Test of Unpredictability of 1-Period Returns):
Significance Level = 0.041

Nonlinear Wald Test (Test that $\delta_t = \delta_t'$):
Significance Level = 0.000

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

$\sigma(\delta_t')/\sigma(\delta_t) = 0.672,$ (0.074) $\quad$ $corr(\delta_t', \delta_t) = 0.175$ (0.146)

$\sigma(\xi_t')/\sigma(\xi_t) = 0.269,$ (0.067) $\quad$ $corr(\xi_t', \xi_t) = 0.915$ (0.064)

Note: figures in parentheses are standard errors.

B. Simulated Results where Efficient Markets Model is True Using Estimated Constrained VAR Model:

1000 Iterations

Linear Wald Test. Rejections at:
5% Level: 0.086, 1% Level: 0.021, Sig. Level Obtained Panel A: 0.072

Nonlinear Wald Test. Rejections at:
5% Level: 0.234, 1% Level: 0.147, Sig. Level Obtained in Panel A: 0.000

$$\delta_t' = \underset{(0.564)}{1.066}\ \delta_t \quad \underset{(0.250)}{-0.034}\ \Delta d_{t-1} \quad \underset{(0.634)}{-0.215}\ \epsilon_t^{30}$$

$\sigma(\delta_t')/\sigma(\delta_t) = 0.926,$ (0.251) $\quad$ $corr(\delta_t', \delta_t) = 0.946$ (0.063)

$\sigma(\xi_t')/\sigma(\xi_t) = 0.878,$ (0.211) $\quad$ $corr(\xi_t', \xi_t) = 0.950$ (0.049)

Note: Ratios and coefficients are means across iterations; figures in parentheses are standard deviations across iterations.

Table 2

Vector Autoregression Results - Time-Varying Discount Rate (Excess Returns)

A. Actual Results (from Campbell and Shiller Table 2 Panel B)

Linear Wald Test (Test of Unpredictability of 1-Period Returns):
Significance Level = 0.028

Nonlinear Wald Test (Test that $\delta_t = \delta_t'$):
Significance Level = 0.000

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

$\sigma(\delta_t')/\sigma(\delta_t) = 0.580$, (0.136) $\quad$ $corr(\delta_t', \delta_t) = 0.309$ (0.341)

$\sigma(\xi_t')/\sigma(\xi_t) = 0.485$, (0.044) $\quad$ $corr(\xi_t', \xi_t) = 0.733$ (0.188)

Note: figures in parentheses are standard errors.

B. Simulated Results where Efficient Markets Model is True Using Estimated Constrained VAR Model:

1000 Iterations

Linear Wald Test. Rejections at:
5% Level: 0.079, 1% Level: 0.019, Sig. Level Obtained Panel A: 0.046

Nonlinear Wald Test. Rejections at:
5% Level: 0.214, 1% Level: 0.133, Sig. Level Obtained in Panel A: 0.009

$$\delta_t' = \underset{(0.571)}{1.045}\,\delta_t \;\underset{(0.254)}{-0.026}\,\Delta d_{t-1} - r_{t-1} \;\underset{(0.641)}{-0.187}\,\epsilon_t^{30}$$

$\sigma(\delta_t')/\sigma(\delta_t) = 0.922$, (0.262) $\quad$ $corr(\delta_t', \delta_t) = 0.948$ (0.065)

$\sigma(\xi_t')/\sigma(\xi_t) = 0.882$, (0.216) $\quad$ $corr(\xi_t', \xi_t) = 0.952$ (0.046)

Note: Ratios and coefficients are means across iterations; figures in parentheses are standard deviations across iterations.

Table 3

Stochastic Simulation - Unit Eigenvalue Imposed

Constant Discount Rate (Real Returns):

1000 Iterations

Linear Wald Test. Rejections at:
5% Level: 0.210, 1% Level: 0.071, Sig. Level of Table 1 Panel A: 0.189

Nonlinear Wald Test. Rejections at:
5% Level: 0.603, 1% Level: 0.493, Sig. Level of Table 1 Panel A: 0.009

$$\delta'_t = \underset{(0.256)}{1.060}\,\delta_t \;\; \underset{(0.141)}{-0.049}\,\Delta d_{t-1} - r_{t-1} \;\; \underset{(0.378)}{-0.459}\,\epsilon^{30}_t$$

$$\sigma(\delta'_t)/\sigma(\delta_t) = \underset{(0.361)}{0.886}, \qquad \mathrm{corr}(\delta'_t,\delta_t) = \underset{(0.297)}{0.533}$$

$$\sigma(\xi'_t)/\sigma(\xi_t) = \underset{(0.212)}{0.658}, \qquad \mathrm{corr}(\xi'_t,\xi_t) = \underset{(0.088)}{0.914}$$

Note: Ratios and coefficients are means across iterations; figures in parentheses are standard deviations across iterations.

Table 4

Stochastic Simulation - Unit Eigenvalue Imposed

Time Varying Discount Rate (Excess Returns)

1000 Iterations

Linear Wald Test. Rejections at:
5% Level: 0.195, 1% Level: 0.071, Sig. Level of Table 2 Panel A: 0.136

Nonlinear Wald Test. Rejections at:
5% Level: 0.589, 1% Level: 0.484, Sig. Level of Table 2 Panel A: 0.134

$$\delta'_t = \underset{(0.246)}{1.043}\,\delta_t \; \underset{(0.122)}{-0.036}\,\Delta d_{t-1} - r_{t-1} \quad \underset{(0.345)}{-0.437}\,\epsilon^{30}_t$$

$\sigma(\delta'_t)/\sigma(\delta_t) = 0.828,$ (0.272)     $\mathrm{corr}(\delta'_t, \delta_t) = 0.533$ (0.318)

$\sigma(\xi'_t)/\sigma(\xi_t) = 0.660,$ (0.206)     $\mathrm{corr}(\xi'_t, \xi_t) = 0.919$ (0.081)

Note: Ratios and coefficients are means across iterations; figures in parentheses are standard deviations across iterations.

## References


Campbell, John Y., and Robert J. Shiller, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," forthcoming, Review of Financial Studies, 1988 (a).

Campbell, John Y., and Robert J. Shiller, "Stock Prices, Earnings, and Expected Dividends," The Journal of Finance, 63, July, 1988 (b).

Fama, Eugene F., and Kenneth R. French, "Permanent and Temporary Components of Stock Prices," Journal of Political Economy, 96:256-73, April, 1988.

Flavin, Marjorie, "Excess Volatility in the Financial Markets: A Reassessment of the Empirical Evidence," Journal of Political Economy, 91:929-56, 1983.

Kleidon, Allan W., "Variance Bounds Tests and Stock Price Valuation Models," Journal of Political Economy, 94:953-1001, October, 1986.

Marsh, Terry A, and Robert C. Merton, "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices," American Economic Review, 76:483-98, June, 1986.

Mattey, Joe, and Richard Meese, "Empirical Assessment of Present Value Relations," Econometric Reviews, 5(2): 171-234, 1986.