



Universidad  
Carlos III de Madrid



This is a postprint version of the following published document:

Banerjee, Anindya ; Dolado, Juan. Do we reject rational expectations models too often?: Interpreting evidence using Nagar expansions. *Economics Letters*, (1987), 24 (1), p. 27-32. Available in [http://www.dx.doi.org/10.1016/0165-1765\(87\)90176-5](http://www.dx.doi.org/10.1016/0165-1765(87)90176-5)

©Elsevier Science Publishers B.V. (North-Holland)

# DO WE REJECT RATIONAL EXPECTATIONS MODELS TOO OFTEN? Interpreting Evidence using Nagar Expansions \*

Anindya BANERJEE

*Jesus College, Oxford OX1 3DW, UK*

Juan DOLADO

*Institute of Economics and Statistics, Oxford OX1 3UL, UK*

## 1. Introduction

In a recent paper, which appeared in this journal, Mankiw and Shapiro (1986) (referred to henceforth as MS) presented Monte Carlo evidence assessing the over-rejection of the orthogonality condition which characterizes rational expectation models. This over-rejection persists even when, asymptotically, the ordinary distribution theory applies to the test. The standard test is to regress the realization of a variable  $X_t$  on the lagged value of some variable  $Y_t$  which belongs to the information set in  $t - 1$ . A classic example is the permanent income hypothesis where  $X_t$  is the change in consumption and  $Y_t$  is income, the idea consisting of testing whether consumption is excessively sensitive to income, once the role of current income in signalling changes in permanent income has been taken into account [see Flavin (1981)].

In general, the canonical model which characterizes the null hypothesis can be described by the following *data generation process* (DGP):

$$X_t = \epsilon_t, \quad Y_t = \lambda Y_{t-1} + \nu_t, \quad E(\epsilon_t, \nu_s) = \delta_{ts} \rho, \quad (1)$$

where  $\epsilon_t$  and  $\nu_t$  are taken to be, without any loss of generality (see MS), white noise processes with unit variance.  $\delta$  represents the Kronecker delta.

Hence the standard procedure is to test the null hypothesis  $H_0: \pi = 0$  in the *model*,

$$X_t = c + \pi Y_{t-1} + u_t, \quad (2)$$

The model is estimated by ordinary least squares.

\* The authors wish to thank Manuel Arellano for his help. Thanks are also due to David Hendry for his encouragement and much useful advice and to members of the Econometrics Workshop at Oxford for their comments. Any remaining errors are attributable solely to the authors.

There are two cases, related to the univariate process that governs the evolution of  $Y_t$ , which are of particular interest. First, if  $\lambda = 1$ , that is,  $Y_t$  follows a random walk, the aforementioned procedure is severely flawed since the two sides of eq. (2) do not have the same order of integration. In such a case, the regression is not consistent and therefore the  $t$ -statistic does not follow the traditional standard distribution [see, for example, MS (1985), Phillips (1986a, b), Phillips and Durlauf (1985), Evans and Savin (1981, 1984) and Banerjee and Dolado (1987)].

The more interesting case and the one that MS (1986) focus on, is where  $Y_t$  follows a borderline stationary process, that is  $|\lambda| < 1$  but close to unity. The model is now 'consistent' since both sides have the same order of integration. However, as shown by MS in their Monte Carlo study, the inferences drawn are still incorrect due to unwarranted reliance on distributions which are correct only asymptotically. The ordinary argument is that the  $t$ -statistic of  $\hat{\pi}$  is likely to have a skewed distribution which is heavily biased downwards from zero in small samples. Inferences based on the use of a symmetric distribution centered on zero are thus likely to be misleading.

This note has two aims. First, we construct a small sample approximation, based on Nagar expansions [see Nagar (1959)], of the expected value of the  $t$ -statistic or more precisely, the continuous normalization of the bias. This does not seem to be available in the existing literature. Second, we use the computed central values to get a feeling for why the critical values reported by MS are so disparate. The approximations, when checked with Monte Carlo estimates, turn out to be quite accurate and shed light on the important role played by the bias, relative to other distributional features, in the results reported by MS.

The paper is organized as follows. Section 2 provides the Nagar expansion for the borderline case of the model appearing in (2) above. Section 3 comments upon the results obtained by using the approximations and compares them with the MS findings. Section 4 concludes the paper.

## 2. Nagar expansions

In order to construct a Nagar expansion of the  $t$ -statistic of  $\hat{\pi}$  in eq. (2), we use the continuous normalization of the OLS estimate of  $\pi$  [see Evans and Savin (1984)]. The normalization factor is derived from the information matrix on the assumption that  $Y_0$  follows the marginal distribution  $N(0, (1 - \lambda^2)^{-1})$ . Then, with small letters denoting deviations from means, we have,

$$ES_T(\lambda)^{1/2} \hat{\pi} = S_T(\lambda)^{1/2} \cdot E \left[ \frac{H}{E(F)[1 - \{1 - F/E(F)\}]} \right] \approx S_T(\lambda)^{1/2} \cdot E \left[ \frac{H}{E(F)} \left[ 2 - \frac{F}{E(F)} \right] \right], \quad (3)$$

$$H = \sum_2^T X_t Y_{t-1}, \quad F = \sum_2^T Y_{t-1}^2, \quad (4)$$

$$S_T(\lambda) = \left[ E \sum_2^T Y_{t-1}^2 - (T-1)^{-1} \left( \sum_2^T E Y_{t-1} \right)^2 \right], \quad (5)$$

where the Nagar expansion makes use of the approximation  $(1+z)^m = 1 + mz + O(m(m-1)z^2)$  for small  $z$ . Hence the error in this approximation has an order of magnitude of  $-2\{F - E(F)\}^2 H / \{E(F)\}^2$ . The following lemma is needed in order to prove the main result of the paper which appears in the theorem below. The proofs are available, upon request, from the authors.

*Lemma.* Let  $\epsilon$  and  $Y_{-1}$  denote, respectively, column vectors of dimension  $(T-1)$  of the form

$$\epsilon = (\epsilon_2, \epsilon_3, \dots, \epsilon_T)', \quad Y_{-1} = (Y_1, Y_2, \dots, Y_{T-1})'.$$

Let  $i$  be a conformable unit vector. Define the following expansions:

$$H = E(Y_{-1}\epsilon'),$$

$$\Omega = E(Y_{-1}Y_{-1}'),$$

$$g = E(\epsilon'Y_{-1}Y_{-1}'\epsilon).$$

Then if (1) characterizes the true DGP with  $|\lambda| < 1$ , with  $Y_0 \sim N(0, (1-\lambda^2)^{-1})$ , the following are true:

$$(a) \quad \text{trace}(\Omega) = (T-1)(1-\lambda^2)^{-1};$$

$$(b) \quad i'Hi = \rho \left[ \frac{T-2}{1-\lambda} - \frac{\lambda(1-\lambda^{T-2})}{(1-\lambda)^2} \right];$$

$$(c) \quad i'H\Omega i = \frac{\rho}{(1-\lambda)^2(1-\lambda^2)^2} \left[ (T-2)(1+\lambda^2)(1-\lambda) + (1-\lambda^{T-2}) \right. \\ \left. \times \{ \lambda^3(T-2) - 3\lambda^2 + \lambda^{T+1} - T\lambda \} + (1-\lambda^2) \right. \\ \left. \times \left[ T\lambda(1-\lambda^{T-2}) + (1-\lambda) - \left[ \frac{1-T\lambda^{T-1} + (T-1)\lambda^T}{1-\lambda} \right] \right] \right];$$

$$(d) \quad i'\Omega i = (1-\lambda^2)^{-1} \left[ 2 \left[ \frac{T-2}{1-\lambda} - \frac{\lambda^2(1-\lambda^{T-2})}{(1-\lambda)^2} \right] - T + 3 \right];$$

$$(e) \quad g = 2\lambda\rho(1-\lambda^2)^{-2} \left[ (T-1) - \frac{(1-\lambda^{2(T-1)})}{(1-\lambda^2)} \right], \quad \text{and}$$

$$(f) \quad i'H\Omega i = i'\Omega Hi.$$

*Theorem.* For the model in (2), if (1) characterizes the true DGP, the corresponding Nagar expansion of the  $t$ -ratio of the  $\pi$  coefficient is as follows:

$$E(t_{\pi=0}) \simeq S_T(\lambda)^{1/2} E(\hat{\pi}) = [E(1) - E(2) + E(3)]/E(4),$$

where

$$E(1) = -(2i'Hi)/\xi^2(T-1)^2;$$

$$E(2) = -(T-1)^2\xi^{-4}[g - 2(T-1)^{-1}i'H\Omega i];$$

$$\begin{aligned}
E(3) &= (T-1)^{-1}\xi^{-4}[(T-1)^{-2}\{\text{trace}(\Omega)\mathbf{i}'\mathbf{H}\mathbf{i} + 2\mathbf{i}'\Omega\mathbf{H}\mathbf{i}\} - 3(T-1)^{-3}(\mathbf{i}'\mathbf{H}\mathbf{i})(\mathbf{i}'\Omega\mathbf{i})]; \\
E(4) &= [(T-1)\xi^2]^{-1/2}; \text{ and} \\
\xi^2 &= (1-\lambda^2)^{-1} - (\mathbf{i}'\Omega\mathbf{i})/(T-1)^2 && \text{if } |\lambda| < 1 \\
&= (T)(T-1)/2 - (\mathbf{i}'\Omega\mathbf{i})(T-1)^{-2} && \text{if } \lambda = 1.
\end{aligned}$$

### 3. Results

To illustrate the working of this approach, we evaluate the expressions in the theorem for values in the parameter space  $(T \times \lambda \times \rho)$  considered by MS, that is,  $T = (50, 200)$ ,  $\lambda = (0.999, 0.99, 0.98, 0.95, 0.90)$  and  $\rho = (1.0, 0.9, 0.8, 0.7, 0.5)$ . The topmost entries of each box in tables 1a and 1b show the computed central values. In all cases we observe that the  $t$ -statistics are centred, as expected, around negative values. Obviously, as  $\lambda$  and  $\rho$  decrease, the expected values converge to the correct ones. It should be noted that in the  $\lambda = 0.999$  case with  $T = 50$ , the approximations proved to be useless, giving quite strikingly large values relative to the others. This implies that for that sample size it would be impossible to separate the 0.999 root from the unit root using our approximations. This is because the order of magnitude of the error term in the Nagar

Table 1a <sup>a</sup>  
Mean  $t$ -ratios and critical values,  $T = 50$ .

$\rho \rightarrow$	1.000	0.900	0.800	0.700	0.500
$\lambda = 0.999$	-	-	-	-	-
	3.000	2.800	2.700	2.500	2.300
	-	-	-	-	-
0.990	-1.171	-1.054	-0.937	-0.820	-0.585
	2.900	2.600	2.500	2.400	2.300
	2.850	2.700	2.600	2.500	2.250
0.980	-1.170	-1.053	-0.935	-0.817	-0.582
	2.800	2.600	2.400	2.300	2.300
	2.850	2.700	2.660	2.500	2.250
0.950	-1.106	-0.995	-0.884	-0.774	-0.552
	2.600	2.400	2.300	2.200	2.100
	2.800	2.650	2.550	2.450	2.200
0.900	-0.952	-0.857	-0.762	-0.667	-0.476
	2.400	2.200	2.200	2.100	2.100
	2.600	2.500	2.450	2.350	2.150

Table 1b <sup>a</sup>  
Mean  $t$ -ratios and critical values,  $T = 200$ .

$\rho \rightarrow$	1.000	0.900	0.800	0.700	0.500
$\lambda = 0.999$	-1.085	-1.066	-0.948	-0.830	-0.593
	2.900	2.800	2.700	2.600	2.400
	2.750	2.700	2.600	2.500	2.250
0.990	-1.066	-0.960	-0.856	-0.746	-0.533
	2.600	2.500	2.400	2.400	2.300
	2.700	2.600	2.500	2.450	2.200
0.980	-1.047	-0.942	-0.840	-0.733	-0.524
	2.500	2.300	2.300	2.200	2.100
	2.700	2.600	2.500	2.400	2.200
0.950	-0.791	-0.712	-0.633	-0.553	-0.396
	2.200	2.200	2.100	2.000	2.000
	2.450	2.350	2.300	2.200	2.050
0.900	-0.574	-0.516	-0.459	-0.401	-0.287
	2.100	2.000	2.000	2.000	2.000
	2.200	2.200	2.100	2.050	1.950

<sup>a</sup> The topmost entry in each grouping of tables 1a and 1b shows the central  $t$ -ratios computed from the Nagar expansions; the middle entry reproduces the Monte Carlo critical values taken from table 1 of Mankiw and Shapiro (1986); the bottom entry shows the critical values obtained, by using the rule described in the text, from the Nagar expansion  $t$ -ratios. For example, the case of  $\rho = 1.00$  and  $\lambda = 0.90$  has a central  $t$ -ratio of  $-0.952$ , an MS critical value of 2.400 and a Nagar expansion critical value of 2.600.

<sup>a</sup> See footnote to table 1a.

Table 2a <sup>a</sup>  
Monte Carlo mean  $t$ -ratios and rejections.  $T = 50$ .

	$\rho = 1.000$	$\rho = 0.700$
$\lambda = 0.990$	-1.362 (0.058) <sup>b</sup> 28(0)	-0.991 (0.064) <sup>b</sup> 15(0)
0.950	-1.201 (0.063) 16(0)	-0.817 (0.065) 9(0)
0.900	-1.034 (0.064) 12(0)	-0.664 (0.067) 8(0)

<sup>a</sup> The Monte Carlo simulations have been carried out using 200 replications in PC NAIVE (copyright D.F. Hendry). The first thirty observations, for each sample size, were discarded. The top entry in each box of tables 2a and 2b shows the mean  $t$ -ratio and the corresponding standard error (in brackets). The bottom entry shows the percentage of rejections using the ordinary two-tailed critical value and the number of rejections in the upper tail using the MS critical values (in brackets).

<sup>b</sup> Denotes significant deviation between the central and the Monte Carlo  $t$ -ratios using a 95% confidence interval.

expansion is no longer negligible. In this fashion, it is possible to establish bounds on the sample size, for a given value of  $\lambda$ , in order to carry out the approximation.

The results indicate quite clearly that there is a strong bias towards over-rejection of the null if conventional asymptotics are used, as exemplified by the MS Monte Carlo results. The next interesting issue is whether we can say anything about the deviations from asymptotic normality of the small sample distribution just using information on the central values of the statistic.

MS report critical values corresponding to the 5% significance level. These are reproduced in the middle entries of each box in table 1. If  $\tau$  is the test statistic, the 5% critical value is the number  $\psi$  such that  $\text{prob}(|\tau| \geq \psi) = 0.05$  under  $H_0$ . As properly noted in footnote 3 of their paper, this is not equivalent to a two-tailed critical value, since the distribution is not symmetric around the origin. Given that the central values are so negative, it might therefore not be too incorrect to infer that all of the rejections occur in the lower tail of the distribution; that is, for large negative values of the  $t$ -ratio, the test is not far from being a one-tailed test at the 5% level. With this in mind, we computed rough approximations of the pseudo two-tailed critical values by adding the one-tailed 5% critical values (1.675 for  $T = 50$  and 1.645 for  $T = 200$ ) to the central values obtained previously. The use of the  $t$ -distribution instead of the asymptotic standardized normal stems from a slight small sample kurtosis correction factor. The results are tabulated in the bottom entries of table 1. When we compare our critical values to those obtained by MS, we observe a close similarity. In the lower range

Table 2b <sup>a</sup>  
 $T = 200$ .

	$\rho = 1.000$	$\rho = 0.700$
$\lambda = 0.990$	-1.172 (0.058) 16(0)	-0.797 (0.066) 9(0)
0.950	-0.802 (0.068) 9(0)	-0.489 (0.070) 7(1)
0.900	-0.582 (0.064) 8(1)	-0.322 (0.069) 6(0)

<sup>a</sup> See footnote <sup>a</sup> of table 2a.

of values for  $\lambda$  and  $\rho$  we seem to be adding slightly more than what is adequate. Obviously for  $\lambda$  and  $\rho$  tending to zero, the critical values converge to 2, since the test becomes a proper two-tailed one and the previous argument does not apply. Replicating *exactly* the MS results, given that the 95% confidence interval for their critical values is  $\pm 0.014$  [variance =  $p(1-p)/N$  where  $N$  = number of replications and  $p = 0.05$ ] would be an unlikely achievement. This notwithstanding, we dare to conclude that our analysis has offered a numerical explanation of the MS results for a reasonable range of borderline cases. The previous evidence also emphasizes the importance of the bias relative to the other features of the distribution, especially those concerning symmetry.

Finally, in tables 2a and 2b, we show the results obtained by a small-scale Monte Carlo study, similar to that carried out by MS, in order to check the accuracy of the approximations and the conjecture about the number of rejections in the upper tail for a representative subset of the parameter space. The evidence on both fronts does not seem to be at odds with the data.

#### 4. Conclusions

Our concern in this paper has been to set up an analytical explanation of recent Monte Carlo results presented by MS on the over-rejection of the random walk property of some rational expectations models when the forcing variable follows a borderline stationary process. Our analysis has served to confirm the statistical pitfalls involved in placing unwarranted confidence in stationary asymptotic results when working with finite samples. It has also emphasized the role of the bias, relative to other distributional features, in explaining the anomalous results. As a by-product we have derived the Nagar expansion of the bias in models of this variety. The analytical results compare favourably with the numerical biases obtained from the Monte Carlo analysis. Current and future research include working out the case in which a trend appears in the regression, as a reflection of the existence of a drift in the autoregressive process, and a refinement of the rough method employed to go from the expected values of the tests to the corresponding critical values.

#### References

- Banerjee, A. and J. Dolado, 1987, Tests of the life cycle-permanent income hypothesis in the presence of random walks: Some evidence using Nagar expansions, Mimeo (Institute of Economics and Statistics, Oxford).
- Evans, G.B.A. and N.E. Savin, 1981, Testing for unit roots: 1, *Econometrica* 49, 1057–1072.
- Evans, G.B.A. and N.E. Savin, 1984, Testing for unit roots: 2, *Econometrica* 52, 1241–1269.
- Flavin, M., 1981, The adjustment of consumption to change in expectations about future incomes, *Journal of Political Economy* 84, 974–1009.
- Mankiw, N.G. and M.D. Shapiro, 1985, Trends, random walks, and tests of the permanent income hypothesis, *Journal of Monetary Economics* 16, 165–174.
- Mankiw, N.G. and M.D. Shapiro, 1986, Do we reject too often?: Small sample properties of tests of rational expectations models, *Economics Letters* 20, 139–145.
- Nagar, A.L., 1959, The bias and moment matrix of the general  $k$ -class estimators of the parameters in simultaneous equations, *Econometrica* 27, 575–595.
- Phillips, P.C.B., 1986a, Time series regression with a unit root, *Econometrica*, forthcoming.
- Phillips, P.C.B., 1986b, Understanding spurious regressions in econometrics, *Journal of Econometrics*, forthcoming.
- Phillips, P.C.B. and S.N. Durlauf, 1985, Multiple time series regression with integrated processes, *Review of Economic Studies* 53, 473–496.