Econometric Reviews
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lecr20

The effect of aggregation on nonlinearity

Clive W. J. Granger [a] & Tae-Hwy Lee [b]

[a] University of California , San Diego, La Jolla, CA, 92093

[b] University of California , Riverside, CA, 92521

Published online: 21 Mar 2007.

To cite this article: Clive W. J. Granger & Tae-Hwy Lee (1999) The effect of aggregation on nonlinearity, Econometric Reviews, 18:3, 259-269, DOI: 10.1080/07474939908800445

To link to this article: http://dx.doi.org/10.1080/07474939908800445

ECONOMETRIC REVIEWS, 18(3), 259–269 (1999)

THE EFFECT OF AGGREGATION ON NONLINEARITY[1]

Clive W.J. Granger and Tae-Hwy Lee

University of California, San Diego, La Jolla, CA 92093
University of California, Riverside, CA 92521

Abstract: This paper investigates the interaction between aggregation and nonlinearity through a monte carlo study. Various tests for neglected nonlinearity are used to compare the power of the tests for different nonlinear models to different levels of aggregation. Three types of aggregation, namely, cross-sectional aggregation, temporal aggregation and systematic sampling are considered. Aggregation is inclined to simplify nonlinearity. The degree to which nonlinearity is reduced depends on the importance of common factor and extent of the aggregation. The effect is larger when the size of common factor is smaller and when the extent of the aggregation is larger.

1. Introduction

Most important macroeconomic series are aggregates, both cross-sectional and temporal. Total consumption may be the aggregate of many million individual consumption figures, and actual consumption decisions may be made more frequently than the available monthly official figures. It might also be claimed that many economic theorists believe that relationships between economic variables are nonlinear and, possibly, it follows that univariate series are generated by nonlinear mechanisms. In this paper the interaction between aggregation and nonlinearity is explored. The basic plan is to consider series generated nonlinearly at the micro level, to then form aggregated series, either cross-sectional or temporal, and then to test these series using various tests for nonlinearity. Thus nonlinearity can be said to be present if it is detected by part of a battery of appropriate tests. The questions that naturally arise are what tests of nonlinearity are appropriate and how powerful they are ? These questions have been investigated in a companion paper by Lee, White, and Granger (1993) - henceforth denoted by LWG. In that paper the results of a large scale simulation using 11 different tests are presented. The tests are based on many different approaches

[1]We wish to thank two referees for useful comments.

including neural network theory, deterministic chaos, Volterra series expansions and various nonlinear parametric forecast models. In general it was found that no single test dominated in all the nonlinear situations considered. In this paper four of the generally better performing tests in LWG are used. These tests are described in Section 3 and their power illustrated in Sections 4 and 5.[2]

As these tests are for linearity it is relevant to ask how this concept should be defined.[3] Suppose that y_t is the series of interest and let X_t be an information set available at time t,

$$X_t : \; y_{t-i}, \; Z_{t-j}, \quad i > 0, \, j \geq 0.$$

For forecasting purposes the second series Z_t, which may be a vector, should only enter the information set with $j > 0$.

Define $g_{n,h} = E(y_{n+h}| X_n)$ being the optimum least squares h-step forecast of y_{n+h} made at time n. $g_{n,h}$ will generally be a nonlinear function of the contents of X_n. Denote $f_{n,h}$ to be the optimum linear forecast of y_{n+h} made at time n, being the best forecast that is constrained to be a linear combination of the contents of X_n. Further, define $e_{n,h} = (y_{n+h} - g_{n,h})$ the h-step forecast error. y_t may be said to be *completely linear* if it obeys both conditions:

(A) $g_{n,h} = f_{n,h}$ all n, h

(B) The conditional distribution of $e_{n,h}$ given X_n is equal to the unconditional distribution of $e_{n,h}$ for all h.

If y_t obeys just condition A for $h = 1$, it can be called *linear in mean.*

Although some tests have power against B being incorrect, especially against forms of heteroskedasticity such as ARCH, in this paper only condition A with $h = 1$ is considered.

2. Some Simple Theoretical Considerations

Consider a bilinear model for a series y_{jt} generated by

$$y_{jt} = \alpha y_{j,t-2}\varepsilon_{j,t-1} + \varepsilon_{jt} \tag{2.1}$$

where $j = 1, \ldots, N$. For example y_{jt} may be the consumption of the jth family at time t. ε_{jt} is a zero mean white noise with decomposition

$$\varepsilon_{jt} = e_t + u_{jt}$$

where e_t, u_{jt} are independent and u_{jt}, u_{kt} are independent for all j, k. If ε_{jt} is viewed as the shock to the jth family, this has a shock component e_t common to all families, the common factor, plus an innovation u_{jt} individual to that family. Substituting into (2.1) and aggregating over j give

$$Sy_t = \alpha e_{t-1} Sy_{t-2} + Ne_t + \alpha \sum_{j=1}^{N} y_{j,t-2}u_{j,t-1} + \sum_{j=1}^{N} u_{jt} \tag{2.2}$$

[2] The power patterns of the tests are similar to those observed in LWG. (There was a programming error for the White's dynamic information matrix test in LWG which is corrected in this paper. We thank Namwon Hyung for pointing it out.) Teräsvirta (1996) explains many of the power results in LWG analytically using the Pitman asymptotic relative efficiency.

[3] See Granger (1998) for more discussion.

where the notation $Sy_t = \sum_{j=1}^{N} y_{jt}$ is used. The last two terms are sums of uncorrelated components and so will have variance $O(N)$ whereas the term Ne_t has variance $O(N^2)$ and so for N large this latter term will dominate. In this case (2.2) will be well approximated by

$$Sy_t = \alpha e_{t-1} Sy_{t-2} + Ne_t \tag{2.3}$$

and so the aggregated series will still follow a bilinear model to a close approximation. However, if $e_t \equiv 0$, so that there is no common factor the first two terms in (2.2) will be absent and there will generally be little or no correlation between terms like $\sum_{j=1}^{N} y_{j,t-2}u_{j,t-1}$ and powers of Sy_{t-2} or products $Sy_{t-2}Su_{t-1}$ and so little or no nonlinearity will remain in the aggregate. This same type of analysis can be extended to other nonlinear models, such as nonlinear autoregressions, a threshold autoregressive and nonlinear relationship between stationary series using various approximations. The suggested results are similar; if common factors are present at the micro level, then some form of nonlinearity is likely to be present in the aggregates but without the common factors nonlinearity is likely to be weak or even nonexistent when aggregation is over large number of components, as pointed out in Granger (1987).

3. The Tests for Nonlinearity

3.1. The Neural Network Test

White (1989) developed a test for neglected nonlinearity likely to have power against a range of alternatives based on a neural network model, a nonlinear flexible functional form being capable of approximating any measurable function. We consider an augmented single hidden layer network in which network output o (a scalar) is determined given input x as

$$o = \tilde{x}'\theta + \sum_{j=1}^{q} \beta_j \psi(\tilde{x}'\gamma_j)$$

where $\tilde{x} = (1\ x')'$; θ is a vector of connection strength from the input layer to the output layer; γ_j is a vector of connection strength from the input layer to the hidden units; β_j is a (scalar) connection strength from the hidden unit j to the output unit; and ψ is a squashing function (e.g., the logistic squasher). Input units $\tilde{x}$ send signals to intermediate hidden units, then each of hidden unit produces an activation ψ that then sends signals toward the output unit. The integer q denotes the number of hidden units added to the affine (linear) network.

When $q = 0$, we have a two layer *affine* network $o = \tilde{x}'\theta$. If the affine network is capable of an exact representation of the unknown function $E(y_t|\,X_t)$, then there exists a vector θ^* such that $H_0 : E(y_t|X_t) = \tilde{X}_t'\theta^*$ with probability one, which constitutes the null hypothesis of interest. This implies $E(e_t^*|X_t) = 0$ where $e_t^* = y_t - \tilde{X}_t'\theta^*$, and thus e_t^* is uncorrelated with any measurable functions of X_t, say $h(X_t)$. That is, $E[h(X_t)e_t^*] = 0$ with probability one. Neural network test is based on a test function h chosen as the activations of "phantom" hidden units $\psi(\tilde{X}_t'\Gamma_j)$, where Γ_j are random column vectors independent of X_t. Thus $E(\Psi_t e_t^*) = 0$ under H_0 where $\Psi_t = (\psi(\tilde{X}_t'\Gamma_1), \ldots, \psi(\tilde{X}_t'\Gamma_q))'$. $E(\Psi_t e_t^*)$ can be estimated by $n^{-1}\sum \Psi_t \hat{e}_t$ where $\hat{e}_t = y_t - \tilde{X}_t'\hat{\theta}_n$ and $\hat{\theta}_n$ is consistent for θ^* when H_0 is true. As in LWG we conduct test on $q^* < q$ principal components of Ψ_t not collinear with X_t (denoted Ψ_t^*) and compute $nR^2 \xrightarrow{d} \chi^2(q^*)$ where R^2 is uncentered squared multiple correlation from a standard linear regression of $\hat{e}_t$ on Ψ_t^*, $\tilde{X}_t$.[4] This test is to determine whether or not augmenting the linear network by including an additional hidden unit with activations would permit an improvement in network performance.

3.2. The Tsay Test

Let y_t be series of interest and denote $X_t = (y_{t-1}, \ldots, y_{t-p})'$ being the information set used to explain y_t. X_t may also include other explanatory variables. Tsay (1986) test is to determine the possibility of improving forecastibility by including product terms such as $y_{t-j}y_{t-k}$ or y_{t-j}^2.

3.3. The White Dynamic Information Matrix Test

White (1987) proposed a specification test based on covariance of conditional score functions. Let y_t be series of interest and X_t the information set. As before, we consider a linear model $y_t = \tilde{X}_t'\theta + e_t$, where $e_t \sim N(0, \sigma^2)$. The log likelihood for this model is

$$\log f_t(X_t, \theta, \sigma) = \text{constant } -\frac{1}{2}\log\sigma^2 - \frac{1}{2\sigma^2}(y_t - \tilde{X}_t'\theta)^2$$

so that, with $u_t = (y_t - \tilde{X}_t'\theta)/\sigma$, the conditional score function is

$$l_t(X_t, \theta, \sigma) \equiv \nabla \log f_t(X_t, \theta, \sigma) = \sigma^{-1}(u_t \;\; u_t X_t' \;\; u_t^2 - 1)'$$

where ∇ is the gradient with respect to θ and σ. Denoting $l_t^* = l_t(X_t, \theta^*, \sigma^*)$, correct specification implies $E(l_t^*) = 0$ and $E(l_t^* l_{t-\tau}^{*\prime}) = 0$, $t = 1, 2, \ldots, \tau = 1, \ldots, t$. Thus we base the test on $m_t = S \operatorname{vec} l_t l_{t-1}'$ where S is a nonstochastic selection matrix focusing attention on particular form of misspecification. Denoting $\hat{l}_t = l_t(X_t, \hat{\theta}_n, \hat{\sigma}_n)$, $\hat{m}_t = S \operatorname{vec} \hat{l}_t \hat{l}_{t-1}'$, and $\hat{\theta}_n$, $\hat{\sigma}_n$ being QMLEs, $n^{-1}\sum_{t=1}^n \hat{m}_t$ should be close to zero under H_0. Then it can be shown that $nR^2 \xrightarrow{d} \chi^2(q)$ under H_0, where R^2 is the squared multiple correlation coefficient from the regression of $\hat{u}_t = (y_t - \tilde{X}_t'\hat{\theta}_n)/\hat{\sigma}_n$ on $\tilde{X}_t$ and $\hat{k}_t$, with $\hat{k}_t$ being defined from $\hat{m}_t = \hat{k}_t \hat{u}_t$, and q is the dimension of m_t.

3.4. The Ramsey RESET Test

From the linear regression $y_t = \tilde{X}_t'\theta + e_t$, let $f_t = \tilde{X}_t'\hat{\theta}$ be the one step linear forecast. Using the polynomials in f_t Ramsey (1969) proposed a test to test $H_0 : c_2 = \ldots = c_k = 0$ in the alternative model $y_t = \tilde{X}_t'\theta + c_2 f_t^2 + \cdots + c_k f_t^k + \nu_t$ for some $k \geq 2$. As in the neural network tests, to retain power without increasing possibility of collinearity, we form the principal components of $(f_t^2, \ldots, f_t^k)$ and regress $\hat{e}_t$ on the $p^* < (k-1)$ largest of them (except the first principal component so as not to be collinear) and $\tilde{X}_t$, which gives an R^2 value. Then nR^2 is distributed as $\chi^2(p^*)$ for n large, under H_0.

3.5. The Simulation Design

For all the simulations, the information set is $X_t = y_{t-1}$ for univariate models, and $X_t = x_t$ for bivariate models.

[4] Teräsvirta, Lin and Granger (1993) studies another test derived from the neural network model.

In performing neural network tests the logistic squasher

$$\psi = [1 + \exp(-\tilde{X}'\gamma)]^{-1}$$

is used with γ being generated from the uniform distribution on [-2, 2] and y_t, X_t being rescaled onto [0, 1]. The number of additional hidden units to the affine network $q = 10$. $q^* = 2$ largest principal components (excluding the first principal component) of these are chosen.

For the White dynamic information matrix tests, appropriate construction of S gives

$$m_t' = \sigma^{-2}(\ u_t u_{t-1} \quad X_t u_t u_{t-1} \quad X_{t-1} u_t u_{t-1} \quad X_t X_{t-1} u_t u_{t-1}\ )$$

so that $q = 4$. In RESET test $k = 5$ and $p^* = 1$ are selected.

5% critical values for the various tests were constructed either using the asymptotic theory or by simulation using a linear AR(1) model $x_t = 0.6x_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0,1)$ for sample size 200 and with 10,000 replications. The empirical power of the four tests at 5% level are computed using 1,000 replications for sample size 200.

4. Effects of Cross-sectional Aggregation

Micro data was generated in two ways, one using a univariate mechanism and the second is a bivariate case. The univariate models were

(a) *Bilinear*

$$y_{jt} = 0.7\, y_{j,t-1}\varepsilon_{j,t-2} + \varepsilon_{jt}$$

(b) *Threshold Autoregressive (TAR)*

$$\begin{aligned} y_{jt} &= 0.9\, y_{j,t-1} + \varepsilon_{jt} \qquad |y_{j,t-1}| \le 1 \\ &= -0.3\, y_{j,t-1} + \varepsilon_{jt} \qquad |y_{j,t-1}| > 1 \end{aligned}$$

(c) *Sign Nonlinear Autoregressive (SGN)*

$$y_{jt} = \mathrm{sgn}(y_{j,t-1}) + \varepsilon_{jt}$$

where $\mathrm{sgn}(x) = 1$ if $x > 0$, 0 if $x = 0$, and -1 if $x < 0$.

(d) *Rational Nonlinear Autoregressive (NAR)*

$$y_{jt} = \frac{0.7\,|y_{j,t-1}|}{|y_{j,t-1}| + 2} + \varepsilon_{jt}.$$

These series were generated for $t = 1, \ldots, n$ and $j = 1, \ldots, m$, so that n is the sample size and m the extent of the aggregation. The values used are $n = 200$, and $m = 1$ (no aggregation) and $m = 20$. The input innovation had three forms, with

$$\mathrm{var}(\varepsilon_{jt}) = 1 \text{ all } j$$

and

$$\varepsilon_{jt} = e_t + \eta_{jt}$$

where e_t, η_{jt} are independent, all j, and η_{jt}, η_{kt} are independent for all k, j and e_t, η_{jt} are always normally distributed with zero mean. The three cases considered are

case (i)	$\text{var}(e_t) = 0$	and	$\text{var}(\eta_{jt}) = 1$
case (ii)	$\text{var}(e_t) = 0.5$	and	$\text{var}(\eta_{jt}) = 0.5$
case (iii)	$\text{var}(e_t) = 0.9$	and	$\text{var}(\eta_{jt}) = 0.1.$

Thus in case (i) there is no common factor and in cases (ii) and (iii) the common factor exists and is of different level of importance.

The bivariate models take the form

$$y_{jt} = g(x_{jt}) + a_{jt}$$

where $x_{jt} = 0.6x_{j,t-1} + \varepsilon_{jt}$; $a_{jt} \sim N(0, \sigma^2)$; a_{jt}, ε_{jt} are independent and ε_{jt} has the three cases as above. Thus any common factor for the y_{jt}'s comes through the x_{jt}. Four values of σ are used, $\sigma = 1, 5, 10, 20$ giving different signal-noise ratios. It is then assumed that both x and y (for their aggregates) are observed, and suitably expanded versions of the tests are used as discussed in the previous section. Two functions $g(x)$ are used, x^2 and $\exp(x)$.

It is assumed that the only quantities observed are $S_m y_t = \sum_{j=1}^{m} y_{jt}$ and equivalently $S_m x_t$, where either $m = 1$ or 20 in the simulation.

Table A1 shows a typical set of results, using the neural network test, 5% critical values and sample size 200. The values shown are the frequencies of times that a null hypothesis of linearity is rejected, out of 1,000 replications, using the simulated critical values (results using the theoretical asymptotical critical values are shown in brackets). The first column shows the case of no aggregation, $m = 1$ and here the common factors are irrelevant, so the figures just illustrate the power of the neural network test against the various nonlinear models. For the four univariate models the power is seen to vary considerably, being low for the nonlinear autoregressive but very high for the sign nonlinear autoregressive models. When applied to the bivariate series, the power is excellent for the higher signal-noise ratios but naturally declined as this ratio goes to smaller. For these two bivariate cases, the signal to noise ratio $\text{var}(g(x))/\sigma^2$ is

σ	1	5	10	20
$\text{var}(x^2)/\sigma^2$	7.0	.28	.07	.019
$\text{var}(\exp(x))/\sigma^2$	2.16	.086	.021	.005

The power of the tests seem to be respectable even with a signal-noise ratio of 2% or less. Results with $\sigma = 1$ and 20 are reported to save space.

The second column shows similar results when aggregation over 20 microunits occurs and there is no common factor. In all cases the tests find nonlinearity less often, as suggested by the theory. The final two columns are with aggregation and different levels of common factor presence. As expected, more nonlinearity is found with aggregation and in bivariate cases even enhances nonlinearity compared to the no aggregation case.

To show that these results do not depend on the test used Tables A2, A3, A4 show the comparable results for the three tests with $n = 200$ and 5% critical values.

5. Effects of Temporal Aggregation

A series may be generated at one time interval but only observed at a greater interval, leading to temporal aggregation of flow variable and systematic sampling of stock

TABLE A.1.
Effects of Cross-Sectional Aggregation
Neural Network Test

Model		case (i) $m = 1$	case (i) $m = 20$	case (ii) $m = 20$	case (iii) $m = 20$
BILINEAR		589	219	402	574
		(575)	(205)	(382)	(562)
TAR		780	61	187	501
		(762)	(53)	(165)	(476)
SGN		982	76	279	780
		(980)	(65)	(260)	(767)
NAR		179	59	112	168
		(156)	(54)	(94)	(151)
SQUARE	$\sigma = 1$	1000	720	1000	1000
		(1000)	(704)	(1000)	(1000)
	$\sigma = 20$	283	64	871	999
		(264)	(58)	(860)	(997)
EXP	$\sigma = 1$	1000	507	1000	1000
		(1000)	(488)	(1000)	(1000)
	$\sigma = 20$	360	70	866	982
		(341)	(67)	(854)	(982)

Power using the simulated critical values is shown. Power using the asymptotic critical values is shown in ( ). Frequencies of rejection out of 1,000 replications are reported at 5% level for sample size 200.

TABLE A.2.
Effects of Cross-Sectional Aggregation
Tsay Test

Model		case (i) $m = 1$	case (i) $m = 20$	case (ii) $m = 20$	case (iii) $m = 20$
BILINEAR		445	169	316	452
		(414)	(146)	(292)	(429)
TAR		52	71	56	68
		(42)	(58)	(43)	(48)
SGN		142	73	72	117
		(122)	(61)	(54)	(98)
NAR		229	75	148	215
		(191)	(55)	(126)	(182)
SQUARE	$\sigma = 1$	1000	824	1000	1000
		(1000)	(796)	(1000)	(1000)
	$\sigma = 20$	370	67	924	1000
		(327)	(47)	(909)	(1000)
EXP	$\sigma = 1$	1000	609	1000	1000
		(1000)	(578)	(1000)	(1000)
	$\sigma = 20$	373	72	889	979
		(343)	(56)	(871)	(978)

TABLE A.3.
Effects of Cross-Sectional Aggregation
Dynamic Information Matrix Test

Model		case (i) $m = 1$	case (i) $m = 20$	case (ii) $m = 20$	case (iii) $m = 20$
BILINEAR		995	561	1000	998
		(995)	(540)	(1000)	(998)
TAR		46	46	71	85
		(41)	(41)	(67)	(76)
SGN		876	102	358	726
		(868)	(92)	(347)	(708)
NAR		114	54	90	121
		(104)	(47)	(86)	(112)
SQUARE	$\sigma = 1$	975	942	989	992
		(973)	(935)	(987)	(992)
	$\sigma = 20$	53	55	116	322
		(50)	(48)	(101)	(306)
EXP	$\sigma = 1$	861	779	960	970
		(856)	(765)	(957)	(969)
	$\sigma = 20$	65	52	157	320
		(60)	(50)	(147)	(302)

TABLE A.4.
Effects of Cross-Sectional Aggregation
RESET Test

Model		case (i) $m = 1$	case (i) $m = 20$	case (ii) $m = 20$	case (iii) $m = 20$
BILINEAR		428	168	309	438
		(408)	(150)	(286)	(416)
TAR		59	70	58	76
		(41)	(57)	(46)	(61)
SGN		369	78	111	235
		(330)	(61)	(84)	(191)
NAR		229	73	147	209
		(188)	(56)	(125)	(187)
SQUARE	$\sigma = 1$	968	820	1000	997
		(967)	(789)	(1000)	(973)
	$\sigma = 20$	190	66	697	773
		(169)	(50)	(672)	(761)
EXP	$\sigma = 1$	632	419	841	665
		(603)	(396)	(834)	(643)
	$\sigma = 20$	206	89	276	368
		(191)	(68)	(248)	(342)

TABLE B.1.
Effects of Temporal Aggregation
Neural Network Test

Model	No aggregation	Systematic $k = 4$	sampling $k = 10$	Temporal $k = 4$	aggregation $k = 10$
BILINEAR	589	310	116	154	43
	(575)	(298)	(107)	(144)	(38)
TAR	780	59	57	53	47
	(762)	(55)	(49)	(49)	(41)
SGN	982	66	43	163	45
	(980)	(60)	(37)	(151)	(40)
NAR	179	58	74	44	43
	(156)	(52)	(61)	(36)	(37)

TABLE B.2.
Effects of Temporal Aggregation
Tsay Test

Model	No aggregation	Systematic $k = 4$	sampling $k = 10$	Temporal $k = 4$	aggregation $k = 10$
BILINEAR	445	222	88	137	47
	(414)	(197)	(74)	(120)	(37)
TAR	52	70	61	57	47
	(42)	(51)	(45)	(44)	(38)
SGN	142	55	55	67	64
	(122)	(38)	(39)	(54)	(43)
NAR	229	66	53	64	58
	(191)	(48)	(42)	(51)	(41)

variable. Systematic sampling occurs when a series x_t is observed at every kth point, giving an aggregate series

$$S_\tau^k = x_{kt}.$$

This may occur with a stock variable, such as temperature, price, money stock, wealth or inventory generated monthly but observed only quarterly, so that $k = 3$. However, some series cannot be measured instantaneously but have to be accumulated over a time period, like flow variables such as rainfall, sales or production. This gives a temporal aggregation where the aggregate is now

$$T_\tau^k = x_{k(t-1)+1} + x_{k(t-1)+2} + \cdots + x_{kt}.$$

TABLE B.3.
Effects of Temporal Aggregation
Dynamic Information Matrix Test

Model	No aggregation	Systematic $k = 4$	sampling $k = 10$	Temporal $k = 4$	aggregation $k = 10$
BILINEAR	995	303	91	133	50
	(995)	(290)	(86)	(128)	(45)
TAR	46	55	43	63	50
	(41)	(54)	(37)	(56)	(44)
SGN	876	123	53	65	53
	(868)	(110)	(51)	(57)	(50)
NAR	114	51	41	49	46
	(104)	(47)	(38)	(47)	(40)

TABLE B.4.
Effects of Temporal Aggregation
RESET Test

Model	No aggregation	Systematic $k = 4$	sampling $k = 10$	Temporal $k = 4$	aggregation $k = 10$
BILINEAR	428	222	87	136	48
	(408)	(199)	(76)	(117)	(38)
TAR	59	70	62	56	47
	(41)	(53)	(46)	(43)	(39)
SGN	369	51	55	67	62
	(330)	(44)	(42)	(53)	(44)
NAR	229	65	53	63	55
	(188)	(50)	(42)	(50)	(43)

The simulation was organized by first generating series using the univariate models discussed in the previous section and then forming temporally aggregated and systematically sampled series of 200 terms after aggregation using $k = 4$ and $k = 10$. Tables B1 to B4 show even when a test is powerful in the no aggregation case, the tests find less evidence of nonlinearity after either temporal aggregation or systematic sampling and generally this effect increases as the extent of the aggregation increases.

The models considered in this study can all be characterized as being 'short memory' in that their optimum forecasts decline to the unconditional mean of the series as the forecast horizon increases. Thus, the expected effect of temporal aggregation on univariate nonlinear series is a decline of any structure, including nonlinearity as the series is sampled less frequently. However, nonlinear relationships between pairs of series can be expected to be less affected by temporal aggregation.

6. Conclusions

Through a simulation study it is found that aggregation is inclined to simplify nonlinearity. If common factors are present at the micro level nonlinearity is likely to remain in the aggregate macro level, but without common factors nonlinearity is likely to decline. It is also seen that nonlinearity is reduced after temporal aggregation or systematic sampling. It is observed that all these are true for all the types of univariate nonlinear series and bivariate nonlinear relationships considered in this paper. The degree to which nonlinearity is declined after aggregation depends on the importance of common factor and the extent of the aggregation. Generally the effect is larger when the size of common factor is smaller and when the extent of the aggregation is larger.

References

Granger, C.W.J. (1987), "Implications of aggregation with common factors," *Econometric Theory,* 3, 208-222.

Granger, C.W.J. (1998), "Overview of nonlinear time series specifications in economics," UCSD, mimeo.

Lee, Tae-Hwy, Halbert White and Clive W.J. Granger (1993), "Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests," *Journal of Econometrics*, 56, 269-290.

Ramsey, J.B. (1969), "Tests for specification errors in classical linear least squares regression analysis," *Journal of Royal Statistical Society*, Series B, 350–371.

Teräsvirta, T. (1996), "Power properties of linearity tests for time series," *Studies in Nonlinear Dynamics and Econometrics, 1, 3-10.*

Teräsvirta, T., Chien-Fu Lin, and C.W.J. Granger (1993), "Power of the neural network linearity test," *Journal of Time Series Analysis*, 14, 209-220.

Tsay, R.S. (1986), "Nonlinearity tests for time series," *Biometrika*, 73, 461-466.

White, H. (1987), "Specification testing in dynamic models," *Advances in Econometrics, Fifth World Congress*, Vol. 1, 1-58, edited by T.F.Bewley, Cambridge University Press.

White, H. (1989), "An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks," *Proceedings of the International Joint Conference on Neural Networks, Washington, D.C.*, San Diego: SOS Printing II: 451-455.