Sensitivity of OLS estimates against ARFIMA error process as small sample Test(s) for long memory.

Anurag N Banerjee
Department of Economics
University of Southampton
Southampton, UK
Email: anb@soton.ac.uk

May 25, 2004

Abstract

Recently there have been much discussion of the theory and applications of long memory processes. In this paper we consider the standard linear model $y = X\beta + u$ and assume that the variance covariance matrix of the errors being generated from an $ARFIMA\,(0, d, 0)$ model. Following Banerjee and Magnus (1999) we investigate the sensitivity of the standard OLS slope (B_L) and sensitivity of variance estimates (D_L) of the linear model near $\theta = 0$. We also investigate the behavior of B_L and D_L under different short memory specifications (for example AR(1) and MA(1) processes) of u. Recalling the Durbin-Watson statistic (DW or $D1$) was related to the sensitivity measure for the OLS variance estimate against ARMA(p,q) errors (Banerjee and Magnus (1999)).This gives us a method to discriminate between long memory and short memory processes, by constucting staistics $B_{L/1}$ and $D_{L/1}$. In this we interpret $D_{L/1}$ as test for long memory process without the shortmemory effects.

1 Introduction

The fractionally integrated autoregressive moving average (ARFIMA) model has recently received considerable attention in economics, but also in other research areas. ARFIMA processes generalize linear ARIMA models by allowing for non-integer differencing powers and do thereby provide a more flexible framework for analyzing time series data. This flexibility enables fractional processes to model stronger data dependence than what is allowed in stationary ARMA models without resorting to non-stationary unit-root processes. However, estimators of the fractional model exhibit larger bias and are computationally more demanding. It is, therefore, beneficial to discriminate fractionally integrated processes from ARMA specifications in a robust modelling step. One way of that is to test the null-hypothesis of an integer differencing power against a fractional alternative. For this purpose the literature frequently utilizes the Geweke and Porter-Hudak (1983) test, the modified rescaled range test of Lo (1991) and Lagrange multiplier tests, see e.g. Agiakloglou and Newbold (1994). The size and power of these asymptotic tests are investigated by Cheung (1993) and Agiakloglou and Newbold. One finding in their studies is the existence of non-negligible small-sample size-distortions.

Some econometricians has also focussed on the Durbin-Watson (DW) test, one of the most intensely studied statistics in all of econometrics as an small sample alternative for testing for long memory process. The properties of a modification of the Durbin-Watson test, due to Nabeya and Tanaka (1990), as a unit root test against short-range dependent alternatives have been studied by Hisamatsu and Maekawa (1994), and against long-range dependent alternatives by Tsay (1998). Nakamura and Tanaguchi (1999) investigate the asymptotics of a standardized Durbin-Watson statistic as a test for independence against fractionally integrated alternatives. They all found that the statistic does well to discriminate between white noise and long-memory alternatives. Since DW statistic was originally designed as a test against AR(1) disturbances, and does well against other short memory alternatives. Therefore it is difficult to distinguish between short memory processes from long memory in finite samples and econometricians hoped that an autocorrelation test will easily detect long memory dependence, perhaps more easily than in the classical AR(1) case.

In this article we consider a different method to discriminate between long-memory and short-memory disturbances. we use a method developed recently by Banerjee and Magnus (1999) who investigated the sensitivity of the OLS estimators from the disturbances' white noise assumption. They considered the standard linear regression model $y = X\beta + u$ under the standard assumptions and that u is normally distributed with mean 0 and covariance matrix is $\sigma^2\Omega(\theta)$, where $\sigma^2 > 0$ and d are unknown. They proposed a pair of sensitivity statistics $B1$ and $D1$ to measure the sensitivity of $\hat{\beta}(\theta)$ and the variance estimator $\hat{\sigma}^2(\theta)$ respectively with respect to changes in the autocorrelation parameter θ of a stationary AR(1) disturbance process. The authors derived the distribution of $B1$, showed that $D1$ has the same form as the Durbin-Watson (DW) test statistic and showed by a simulation experiment that $B1$ and $D1$ are nearly independent. Their main conclusions are i) the predictor is not sensitive to covariance misspecification, but $\hat{\sigma}^2(\theta)$ can be very sensitive. ii) the statistic $B1$ and D1 can still be used for general ARMA(p,q) disturbances.

In this paper we are interested in the sensitivity of $\hat{\beta}(d)$ and the variance estimator $\hat{\sigma}^2(d)$ measured by the statistic B_L and D_L respectively when the error process u is distributed $ARFIMA(d)$. We show that the properties B_L and D_L are different from $D1$ and $B1$ which were developed for short memory error processes. Further we shall shall propose a sensitivity statistics $B_{L/1}$ and $D_{L/1}$ to distinguish between $ARFIMA(d)$ process and short- Memory processes.

2 Preliminaries

We consider the standard linear regression model

$$y = X\beta + u, \tag{2.1}$$

where y is an $T \times 1$ random vector of observations, X a non-random $T \times k$ matrix of regressors, β a $k \times 1$ vector of unknown parameters and u an $T \times 1$ vector of random disturbances. We assume that X has full column-rank k. The stochastic component u_t follows an Long Memory process, with normal innovations,

$$\begin{aligned} (1-L)^d u_t &= \varepsilon_t \quad t = 0, 1 \ldots, T \\ u_t &= 0, \;\; t < 0. \end{aligned} \tag{2.2}$$

with $\varepsilon_1, \ldots, \varepsilon_T \sim iid\ N(0, \sigma^2)$, where $\sigma^2 > 0$.Formally, for any real d,

$$(1-z)^{-d} = \sum_{h=0}^{\infty} \frac{\Gamma(h+d)}{\Gamma(h+1)\Gamma(d)} z^j$$

with $\Gamma(.)$ denoting the Gamma Function. This leads to a autocovariance function (Beran (1994)) of

$$\gamma_h(d) = \sigma^2 \frac{(-1)^h \Gamma(1-2d)}{\Gamma(h-d+1)\Gamma(1-h-d)}.$$

The correlations are equal to

$$\rho_h(d) = \frac{\Gamma(1-d)\Gamma(h+d)}{\Gamma(d)\Gamma(h+1-d)} \tag{2.3}$$

Therefore using (2.3) we can write the variance matrix of u as,

$$\Omega(d) = \sigma_d^2 \sum_{h=0}^{T-1} \rho_h(d) T^{(h)} \tag{2.4}$$

where we denote by $T^{(h)}, 0 \le h \le T-1$, the $T \times T$ symmetric Toeplitz matrix with

$$T^{(h)}(i,j) = \begin{cases} 1 & \text{if } |i-j| = h, \\ 0 & \text{otherwise.} \end{cases}.$$

Differentiating both sides of (2.4) with respect to d at $d = 0$, then yields

$$A_L = \frac{\partial \Omega(d)}{\partial d} \mid_{d=0} = \sum_{h=1}^{T-1} \frac{1}{h} T^{(h)}. \tag{2.5}$$

3 Sensitivity of the predictor and the variance estimate

If d is known, then the parameters β and σ^2 can be estimated by generalized least squares. Thus,

$$\hat{\beta}(d) = (X'\Omega(d)^{-1}X)^{-1}X'\Omega(d)^{-1}y \tag{3.1}$$

and

$$\hat{\sigma}^2(d) = \frac{(y-\hat{y}(d))'\Omega(d)^{-1}(y-\hat{y}(d))}{T-k}, \tag{3.2}$$

where $\hat{y}(d)$ denotes the predictor for y, that is,

$$\hat{y}(d) = X\hat{\beta}(d). \tag{3.3}$$

We wish to assess how sensitive (linear combinations of) $\hat{\beta}(d)$ are with respect to small changes in d when d is close to 0. The predictor is the linear combination most suitable for our analysis. Since any estimable linear combination of $\hat{\beta}(d)$ is a linear combination of $\hat{y}(d)$, and vice versa, this constitutes no loss of generality.

Definition 1 *We define the sensitivity of the predictor $\hat{y}(d)$ (with respect to d) as*

$$z_L = \left.\frac{\partial \hat{y}(d)}{\partial d}\right|_{d=0}. \tag{3.4}$$

The sensitivity of $\hat{\beta}(d)$ (with respect to d) is then

$$\left.\frac{\partial \hat{\beta}(d)}{\partial d}\right|_{d=0} = (X'X)^{-1}X'z_L.$$

In order to use the (normally distributed) $T \times 1$ vector z_L as a sensitivity statistic, we transform it into a quadratic form in the usual way. We thus propose

$$B_L = \frac{z_L'(C_L C_L')^- z_L}{(T-k)\hat{\sigma}^2(0)}, \tag{3.5}$$

as a statistic to measure the sensitivity of the predictor $\hat{y}(d)$ with respect to d. (The notation A^- denotes a generalized inverse of A.), where $M = I_T - X(X'X)^{-1}X'$ be the usual idempotent matrix and

$$C_L = (I_T - M)A_L M \tag{3.6}$$

Large values of B indicate that $\hat{y}(d)$ is sensitive to small changes in d when d is close to 0 and therefore that setting $d = 0$ is not justified. The statistic B_L depends only on y and X and can therefore be observed. Since the distribution of y depends on d, so does the distribution of B. We now state our main result.

In order to assess the sensitivity of the variance estimator $\hat{\sigma}^2(d)$ with respect to small changes in d.

Definition 2 *We define the sensitivity of $\hat{\sigma}^2(d)$ (with respect to d) as*

$$\lambda_L = \left.\frac{\partial \hat{\sigma}^2(d)}{\partial d}\right|_{d=0}. \tag{3.7}$$

Upon scaling we find

$$D_L = \frac{\lambda_L}{\hat{\sigma}^2(0)} = \left.\frac{\partial \ln \hat{\sigma}^2(d)}{\partial d}\right|_{d=0} \tag{3.8}$$

as a suitable statistic for our purpose.

Given that the sensitivity measures exist, and since we are interested in knowing how close are $\hat{\sigma}^2(d)$ and $\hat{\sigma}^2(0)$ (or for the matter $\hat{\beta}(d)$ and $\hat{\beta}(0)$) we need to study how close D_L (or B_L) is to zero. The D_L (or B_L) will generally be random variables[1]. So it will be useful to study the following probabilities as a measure of "closeness" to zero,

$$\pi\left(d,:S\right)=\Pr_d\left(|S|\geq z_\alpha\right),\, S=D_L(orB_L), \tag{3.9}$$

where $\Pr_d$ is the probability measure associated with the random variable $u\sim N\left(0,\sigma^2\Omega\left(d\right)\right)$, and z_α is obtained from the equation,

$$\Pr_0\left(|S|\geq z_\alpha\right)=\alpha\ ,\ 0<\alpha<1. \tag{3.10}$$

where $\Pr_0$ is the probability measure associated with white noise. $\pi\left(d\right)$is essentially a robustness function of the D_L (or B_L) statistic, against long memory error process. The probabilities give an indication of how close to zero the sensitivity measures are. The greater the probability mass of the sensitivity measures around zero, closer is the distance between $\hat{\sigma}^2(d)$ and $\hat{\sigma}^2(0)$ (or $\hat{\beta}(d)$ and $\hat{\beta}(0)$). In order to have a sharper bound for the sensitivity we will choose a lower value of α (in this paper we chose $\alpha=0.05$). Higher the value of $\pi\left(d\right)$, higher is the probability of sensitivity of D_L (or B_L). In this sense the sensitivity of the relevant statistic increases when $\pi\left(d\right)$ increases. The statistic D_L can be seen as an long memory equivalent of the Durbin-Watson statistic.

Theorem 3 *We have*

1. $z_L=-C_Ly$;
2. $B_L=\frac{y'W_Ly}{y'My},\quad W_L=C_L'(C_LC_L')^{-}C_L$;
3. *If* $0<r_L<T-k$ *and the distribution of* y *is evaluated at* $d=0$, *then*

$$B_L\sim Beta(r_L/2,(T-k-r_L)/2).$$

Theorem 4 *We have*

(a) $\lambda_L=-\frac{y'MA_LMy}{T-k}$;

(b) $D_L=-\frac{y'MA_LMy}{y'My}$;

(c) *If the distribution of* y *is evaluated at* $d=0$, *then*
$D_L=-\frac{v'P'A_LPv}{v'v}$,

where P *is an* $T\times(T-k)$ *matrix containing the* $T-k$ *eigenvectors of* M *associated with the eigenvalue 1, that is,* $M=PP',P'P=I_{T-k}$, *and* $v\sim N(0,I_{T-k})$.

[1]See Section 5 where $\lambda_l^{(1)}\left(A\right)=0$, $A=A1,A2$, when the deterministic component is linear $\left(d_t=\beta_1+t\beta_2\right)$.

Using Pitman's Lemma we can obtain the moments of D_L exactly.

Lemma 5 *(Pitman (1937), Laha (1954)). Let $x_1, \ldots, x_T$ be identically and independently distributed random variables with a finite second moment. Then $\sum_i a_i x_i / \sum_i x_i$ and $\sum_i x_i$ are independent if and only if each x_i follows a gamma distribution.*

A consequence of Pitman's Lemma is that when $u \sim N(0, I_T)$, then

$$\frac{u'MA_LMu}{u'u} \text{ is independent of } u'u. \tag{3.11}$$

which leads to

$$E\left(\frac{u'MA_LMu}{u'u}\right)^s = \frac{E\left(u'MA_LMu\right)^s}{E\left(u'u\right)^s}. \tag{3.12}$$

Hence

Theorem 6 *If $u \sim N(0, I_T)$, then*

1.
$$E\left(D_L\right) = 0 \text{ and } var\left(D_L\right) = trace\left(A_L M\right)^2,$$

2. *further*
$$\lim_{T\to\infty} var\left(D_L\right) \leq \frac{\pi^2}{6}$$
with the equality holding when if $X = 0$.

Proof. (later) ■

We known from Theorem 3 that B_L follows a Beta distribution and 4 and 4 gives us the properties of D_L when the disturbances are white noise. The logical next step is to ask how B_L and D_L behaves when the disturbances follows ARFIMA(0,d,0) long memory stationary process.

We have 10 data sets; (see Table 1 for details) . For each dataset we calculate $\pi\left(d : B_L\right)$ and $\pi\left(d : D_L\right)$ such that $\alpha = 0.05$ under the assumption that the disturbances are ARFIMA(0,d,0) for values of d between 0 and 1/2. As noted before, the D_L-statistic is essentially the long memory equivalent of the short memory DWstatistic. As a result, $\pi\left(d : D_L\right)$ can be interpreted as the power of D_L in testing $d = 0$ against $d > 0$. Alternatively we can interpret $\pi\left(d : D_L\right)$ as the sensitivity of $\hat{\sigma}^2$ with respect to d. In the same way, B_L measures the sensitivity of $\hat{y}$ (and $\hat{\beta}$) with respect to d. One glance at Figure 2 shows that B_L is quite moderately sensitive, respect to d. This is in contrast to the results of the sensitivity analysis of the $B1$ statistic obtained in (Banerjee and Magnus 1999), where they show that the $\hat{y}$ (and $\hat{\beta}$) in most cases are insensitive to ARMA type processes. The figure for D_L also shows high sensitivity of the variance estimator $\hat{\sigma}^2$. The figures shows the probabilities for $n = 100$ The main conclusion is that not only D_L is sensitive to d but also B_L shows moderate sensitivity in around 30-40% of the cases.

Figure 1

Figure 2

We shall also investigate the question of how B_L and D_L when the disturbances follow a stationary AR(1) and MA(1) process. For each dataset in Figure 2 and 3 we have calculated

$$\pi(\theta : S) = \Pr_{\theta}(|S| \geq z_{\alpha}), \; S = D_L(or B_L), \tag{3.13}$$

where $\Pr_{\theta}$ is the probability measure assuming the error disturbances are distributed as an AR(1) process with parameter θ.for values of θ between 0 and 0.5.

Figure 3

Figure 4

We see from Figure 3 and Figure 4 that B_L shows no sensitivity against AR(1) process and D_L shows only a moderate amount of sensitivity against the short memory AR(1) process.

Similarly in Figure 4 and 5 we calculate

$$\pi(\phi : S) = \Pr_{\phi}(|S| \geq z_{\alpha}), \; S = D_L(or B_L), \tag{3.14}$$

under the assumption that the disturbances are MA(1) with parameter ϕ.for values of ϕ between 0 and 0.5.

Figure 5

Figure 6

From Figure 5 and Figure 6 we see that the B_L statistic shows no sensitivity against MA(1) disturbances.[2]

In the next section we shall try to devise a sensitivity statistic which can discriminate between Long memory and short memory processes. This implies we need a statistic which would show more sensitivity toward long-memory process than short memory process.

4 Long Memory Sensitivity without Short Memory

The first step away from white noise disturbances is an AR(1) process or in general the ARMA(p,q) processes, which are in general classified as short-memory process since the decay of the autocorrelation functions decay faster than the ARFIMA(0,d,0) processes. In a seminal paper, Hosking (1984) observes that a long-memory process can be approximated by an ARMA(1,1) process reasonably well when the approximating ARMA process has both roots close to the unit circle. Although no rigorous justification of this assertion is given in his paper,simulation studies conducted in Hosking (1984) indicate the validity of this assertion. Recently a paper by Basak, Chan and Palma (2001) proposed a mean square error criterion based approximation a long-memory time series by a short-memory

[2] The probabilities were all calculated using our own adaptation of Imhof's (1961) routine in Gauss which is availablevia internet under http://www.american.deu/academic.depts/cas/econ/gaussres/Gausidx.htm.

ARMA(1, 1) process. Here we shall use or sensitivity measures to see whether we can discriminate between short-memory and long memory processes.

In order to measure the sensitivity of $\hat{\sigma}^2(\theta)$ the variance estimator $\hat{\beta}(\theta)$ the slope estimator respectively with respect to changes in the autocorrelation parameter, θ of a stationary AR(1) (or MA(1) or indeed ARMA(1,1)) disturbance process, Banerjee and Magnus (1999) proposed a pair of sensitivity statistics

$$D1 = \frac{\hat{u}'T^{(1)}\hat{u}}{\hat{u}'\hat{u}} \tag{4.15}$$

and

$$B1 = \frac{\hat{u}'W^{(1)}\hat{u}}{\hat{u}'\hat{u}} \quad \text{such that } W^{(1)} = C^{(1)\prime}(C^{(1)}C^{(1)\prime})^{-}C^{(1)} \tag{4.16}$$

where $C^{(1)} = (I_T - M)T^{(1)}M$, respectively The authors derived the distribution of $B1$ which is similar to Theorem 3. They also showed that $D1$ has the same form as the Durbin-Watson (DW) test statistic rather the alternative DW statistic proposed by King (1981).[3]

The asymptotic properties of DW statistic under long memory processes was studied by Tsay. (1998), Nakamura and Taniguchi (1999). A recent paper by Kleiber and Kramer (2004) studies the finite sample properties of DW under ARFIMA(0,d,0). These studies conclude that the DW statistic is a powerful test. This implies that the D1 by (4.15.) is highly sensitive to long memory process.

So the question remains can we distinguish between short-memory and long memory processes using our sensitivity measures B1, B_L,D1 and D_L. One obvious way is to distinguish to take the effect of the short memory process out of the long memory process by subtracting the relevant statistics.

Definition 7 *We define the sensitivity of $\hat{\sigma}^2(d)$ (with respect to d, the long memory parameter) without the short memory effect as*

$$D_{L/1} = D_L - D_1$$

Definition 8 *We define the sensitivity of the predictor $\hat{y}(d)$ (with respect to d,the long memory parameter) without the short memory effect as*

$$B_{L/1} = B_L - B_1$$

We can therefore write

$$D_{L/1} = \frac{\hat{u}'A_{L/1}\hat{u}}{\hat{u}'\hat{u}}$$

where $A_{L/1} = A_L - T^{(1)}$, and

$$B_{L/1} = \frac{\hat{u}'W_{L/1}\hat{u}}{\hat{u}'\hat{u}}$$

where $W_{L/1} = W_L - W^{(1)}$.

Theorem 9 *If $u \sim N(0, I_T)$, then*

[3] See Appendix for the statements of the theorems.

1.

$$Corr\left(D_{L/1}, D1\right) = \frac{trace(D_{L/1}PD_1P)}{n-k}$$

where $P = I - M$,

2. further

$$Cov\left(D_{L/1}, D1\right) = 0$$

when $X = 0$.

Theorem 3, implies that $D_{L/1}$ and $D1$ are nearly independent statistics. Therefore the information contained in $D_{L/1}$ is independent of the outcome of the $D1$ (DW tests). Therefore we can think $D_{L/1}$ as a way of measuring the pure long-memory content of the error process.

As before we shall also investigate the question of how $B_{L/1}$ behave when the disturbances follow a stationary are ARFIMA(0,d,0), AR(1) and MA(1) process. To do that we have calculated $\pi\left(d:S\right), \pi\left(\theta:S\right)$ and $\pi\left(\phi:S\right)$ where $S = B_{L/1}$ or $D_{L/1}$.

Figure 7

Figure 8

Figure 9

Figures 7 -9 plots the sensitivity of the $B_{L/1}$ statistic. and

Figure 10

Figure 11

Figure 12

Figures 10- 12 plots the sensitivity of the $D_{L/1}$ statistic when the error disturbances are ARFIMA(0,d,0), AR(1) and MA(1) respectively. One interesting fact is that the $D_{L/1}$ statistic shows no sensitivity to short memory processes (AR(1) and MA(1))[4] but highly sensitive to the long memory process (ARFIMA(0,d,0) which implies that the $D_{L/1}$ statistic can distinguish between long memory and short memory processes. On the other hand when we look at the sensitivity statistic of the predictor namely the $B_{L/1}$ statistic it fails to distinguish between AR(1) and Long memory processes, though the $B_{L/1}$ shows insensitivity under MA(1) process.

Recall $\pi\left(\theta:D1\right)$ can be interpreted as the power of $D1$ or DW in testing $\theta = 0$ against $\theta > 0$. Similarly we can interpret $\pi\left(.:D_{L/1}\right)$ as a power curve for testing for long memory against short-memory alternatives.

[4]This is in accordance with King and Evans (1988).

5 Conclusion

In this article we have introduced a new sensitivity measures, B_L and D_L, which is designed to decide whether the predictor and the variance estimators are sensitive long memory (in particular ARFIMA(0,d,0)) misspecification. Our results show that the OLS estimator $\hat{\beta}$ (or predictor $\hat{y}$) are is moderately sensitive ARFIMA(0,d,0) misspecification, which is in contrast to the results of Banerjee and Magnus (1999) where they conclude that the OLS slope estimator is robust to short memory ARMA specification. The D_L statistic which measures the sensitivity of the variance as expected shows non-robustness ARFIMA(0,d,0) and moderately robust against AR(1) disturbances, but shows very little sensitivity against MA(1) disturbances.

We then device a sensitivity measures $B_{L/1}$ and $D_{L/1}$ for the predictor and the variance estimators which removes the short-memory effects and purely measures the long memory misspecification. And indeed we find that $D_{L/1}$ statistic shows no sensitivity to short memory processes but is highly sensitive to the long memory process . Unfortunately, $B_{L/1}$ statistic it fails to distinguish between short memory and Long memory processes.

Therefore we conclude that $D_{L/1}$ will be a useful statistical measure to distinguish between long memory and short-memory processes.

6 References

Agiakloglou, C. and P. Newbold (1994) Lagrange Multiplier Tests for Fractional Difference. Journal of Time Series Analysis 15 (3), 253-262.

Banerjee, A. and J.R. Magnus (1999), The sensitivity of OLS when the variance matrix is (partially) know, Journal of Econometrics 92, 295-323.

Basak, G.P., Chan, N.H., and Palma, W. (2001). Approximation of a long-memory time series with an ARMA(1,1) model, Journal of Forecasting 20, 367-389.

Beran, J., (1994) Statistics for long-memory processes, (Chapman & Hall), New York.

Cheung, Y.-W. (1993) Tests for Fractional Integration: A Monte Carlo Investigation. Journal of Time Series Analysis 14 (4), 331-345.

Geweke, J. and S. Porter-Hudak (1983) The Estimation and Application of Long Memory Time Series Models. Journal of Time Series Analysis 4
(4), 221-238.

Hisamatsu, H., and Maekawa, K. (1994). The distribution of the Durbin-Watson statistic in integrated and near-integrated models. Journal of Econometrics, 61, 367-382.

Hosking J.R.M. (1984),. Modelling persistence in hydrological time series using fractional differencing. Water Resources Research 20: 1898–1908.

Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. Biometrika, 48, 419-426.

King. M.L., and Evans, M.A. (1988). Locally optimal properties of the Durbin-Watson test, Econometric Theory, 4, 509-516.

Kleiber, C. and Walter Kr¨amer, (2004) Finite sample power of the Durbin-Watson test against fractionally integrated disturbances, Working Paper, Fachbereich Statistik, Universit¨at Dortmund, Germany.

Lo, A. W. (1991) Long Term Memory in Stock Market Prices. Econometrica 59 (5), 1279-1313.

Mathai, A.M. and S.B. Provost, (1992), Quadratic Forms in Random Variables: Theory and Applications (Marcel Dekker), New York.

Nakamura, S., and Taniguchi, M. (1999). Asymptotic theory for the Durbin-Watson statistic under long-memory dependence. Econometric Theory, 15, 847-866.

Tsay, W.J. (1998), "On the Power of Durbin-Watson Statistic against Fractionally Integrated Processes," Econometric Reviews, 17, 361-386.

7 Appendix:

7.1 Tables

X	**Regressors (T=100)**
1	Constant,Time Trend,Random Numbers $\sim N(0,1)$
2	Constant,Time Trend,Random Numbers $\sim Beta\left(\frac{1}{2},\frac{1}{2}\right)$
3	Constant,Time Trend,Random Numbers $\sim Gamma(2)$
4	Constant,Random Numbers $\sim N(0,1)$,Random Numbers $\sim Beta\left(\frac{1}{2},\frac{1}{2}\right)$
5	Random Numbers $\sim Beta\left(\frac{1}{2},\frac{1}{2}\right)$,Random Numbers $\sim Gamma(2)$, Dummy variable $\sim I[U(0,1) >$
6	Time Trend,Random Numbers $\sim N(0,1)$,Dummy variable generated by $I[U(0,1) > 0.5]$
7	Time Trend,Random Numbers $\sim N(0,1)$,Random Numbers $\sim Gamma(2)$
8	Time Trend,Random Numbers $\sim N(0,1)$,Random Numbers $\sim Beta\left(\frac{1}{2},\frac{1}{2}\right)$
9	Random Numbers $\sim N(0,1)$,Random Numbers $\sim Beta\left(\frac{1}{2},\frac{1}{2}\right)$,Random Numbers $\sim Gamma(2)$
10	Time Trend,Random Numbers $\sim Beta\left(\frac{1}{2},\frac{1}{2}\right)$,Random Numbers $\sim Gamma(2)$

7.2 Figures:

Figure 1

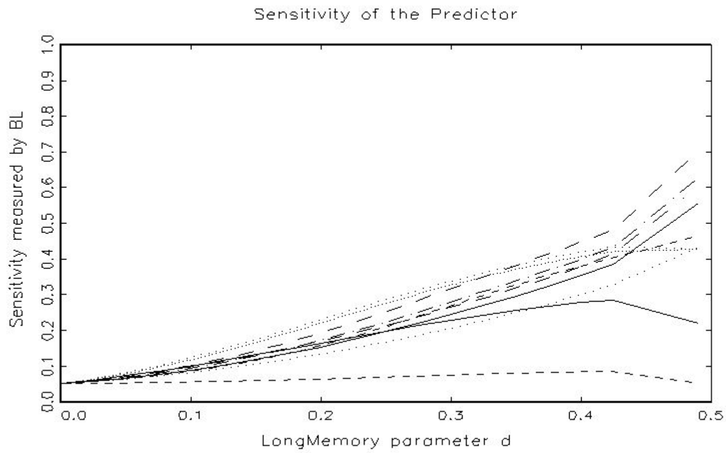

Figure 2

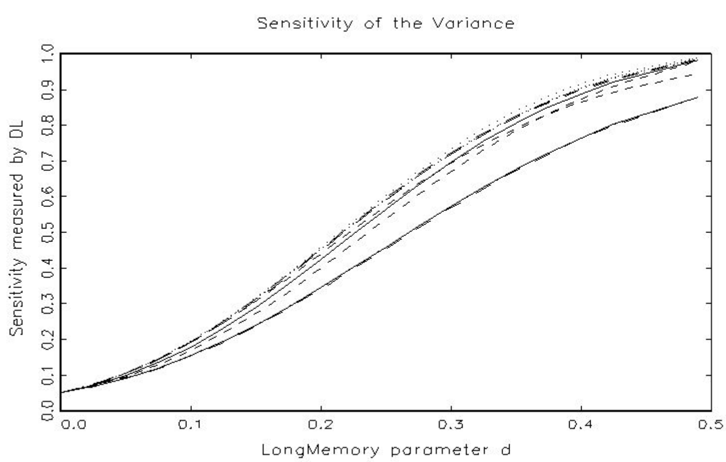

Figure 3

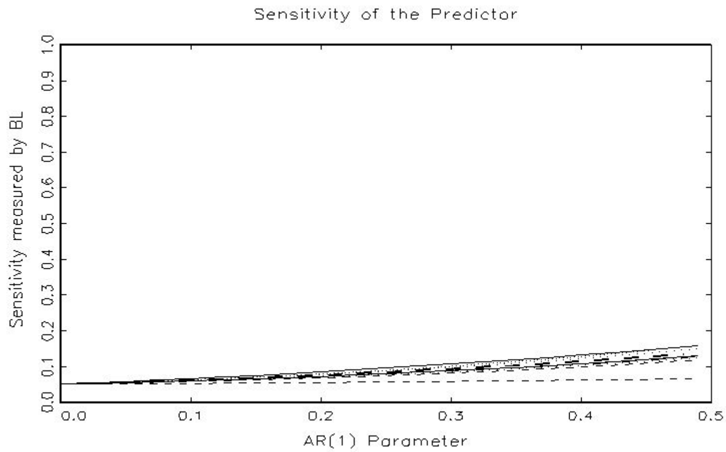

Figure 4

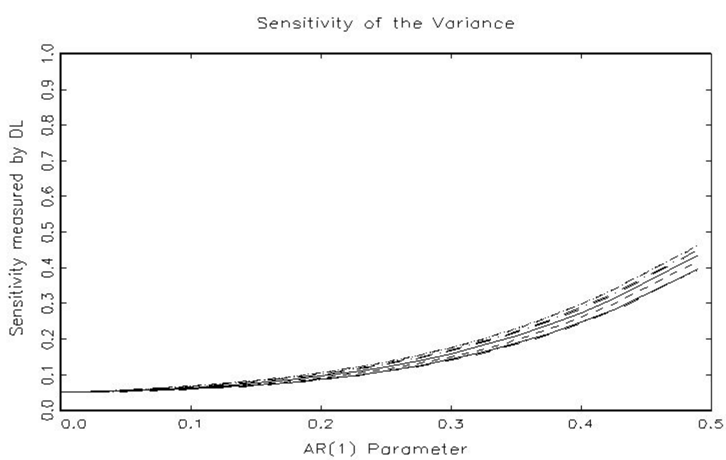

Figure 5

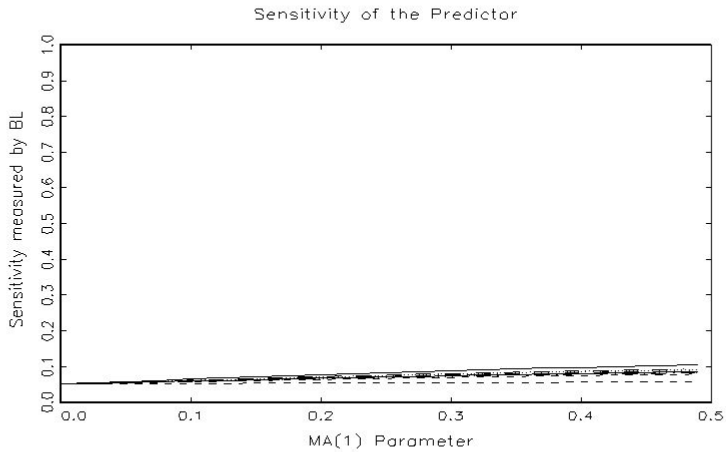

Figure 6

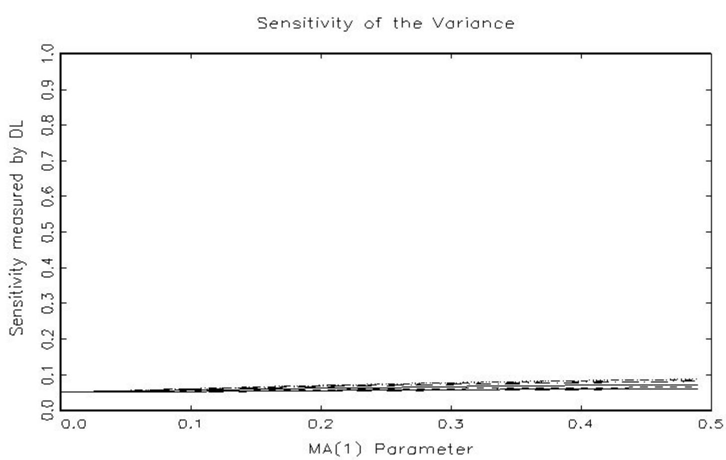

Figure 7

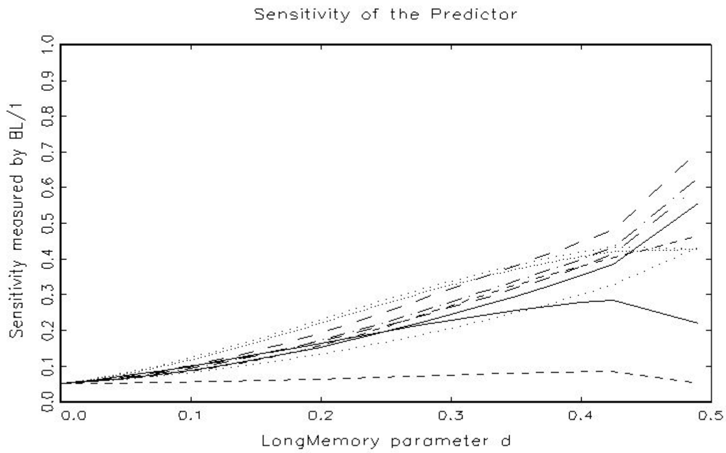

Figure 8

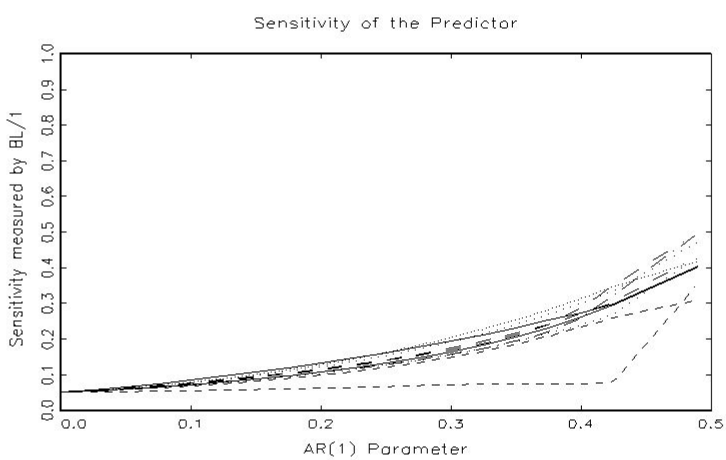

Figure 9

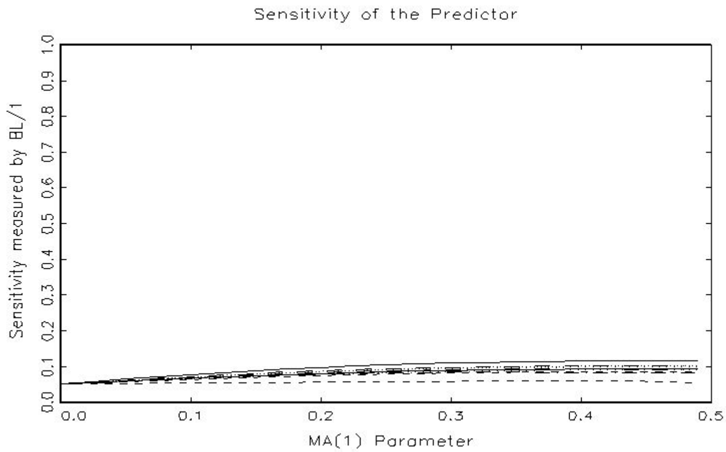

Figure 10

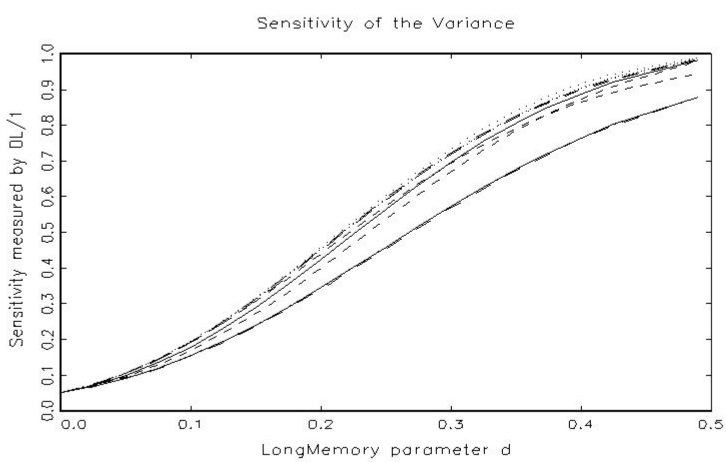

Figure 11

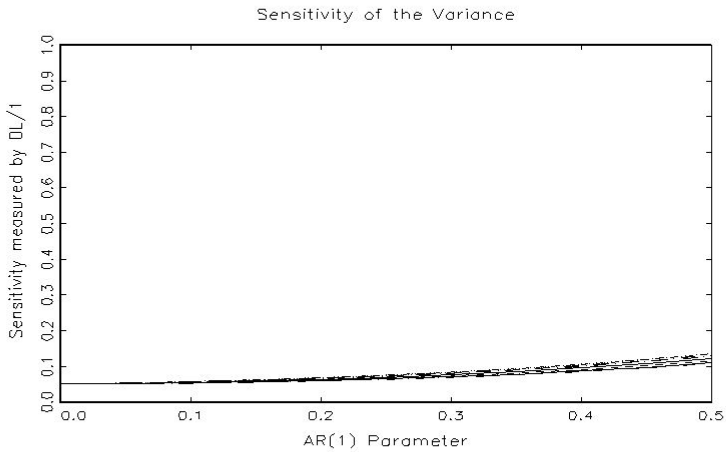

Figure 12