Estimation of the Long Memory Parameter in the Presence of Additive Outliers and Inliers: Some Simulation Evidence

by

Mohammad Ali Ashraf

Major paper presented to the Department of Economics of the University of Ottawa in partial fulfillment of the requeriments of the M. A. Degree

Supervisor

Professor Gabriel Rodríguez

ECO 7997

Ottawa, Canada

May 2001

Acknowledgements

This is my greatest pleasure to express reverence and to acknowledge the pivotal assistance of my research supervisor, Professor Gabriel Rodriguez, Department of Economics of the University of Ottawa, Canada, for providing such a wonderful congenial environment in which to undertake this research endeavor. As a matter of fact, I was rendered all kinds of supportive tips from him in designing the outline of the study, selecting data-set, modeling methodology and finally presenting the paper in the most elegant fashion. I feel no hesitation to reveal here that it was absolutely impossible for me to accomplish such an accomplishment in the field of econometric research without his skillful guidance..

An expression of sincere gratitude is particularly due to Professor Kathleen Day for her constructive comments in every details and helpful suggestions. She has also provided me an editorial support which was mostly needed for presenting my paper. I must mention here that without her patience and unstinting cooperation, this study may not be completed in due date.

I must recall my previous University teacher, Professor Atiur Rahman Mollah, Department of Agricultural Economics, Bangladesh Agricultural University (BAU), Mymensingh, Bangladesh, who first provided me an opportunity of doing research in economic field under his direct supervision during his tenure as the Project Director of a Research Project going on in the Bureau of Socio-Economic Research and Training (BSERT), an affiliated research station of BAU.

It would be a great mistake on my part if I don't mention the name of my another respectable teacher, Mr. Bhola Nath Shaha Chowdhury, who taught me an ABC of English language in a High School in Bangladesh. He was the man who instilled a lot of moral ethoes into the mind at that very tender age of my life.

I would also like to remember my honorable parents, especially my departed father —-my mentor, who enlightened my life with the spectrum of his life-long experience and wisdom which showed me the path for making my life enshrined with the eternal light of knowledge. Nevertheless, a particular respect is due to my only elder brother, Dr. Mohammad A. Akbar and his wife, Dr. Afroza K. Mili for providing me all kinds of support and constant encouragement to go for higher studies. Towering above of these forms of assistance is the mental support that I received from my beloved wife Arifa Yasmeen Ashrafee.

Last but not least, my special thanks and high esteem go to Luce Laviolette, the administrative Assistant and the Secretarial staff for the help and cooperation they provided me throughout the entire program. However, the errors may have remained are absolutely mine.

Abstract

In this paper, we analyze the effects on the estimation of the fractional parameter, d, of additive outliers or additive inliers. There are many methods of estimating the fractional parameter but here, we consider only the method of estimation proposed by Geweke and Porter-Hudak (GPH) (1983), which is a method based on the frequency domain. The paper includes a brief review of the origin of long memory models, a description of the different approaches and the building block models proposed in the literature. We also discuss the GPH estimation method and present a brief survey of the more important empirical applications of this kind of model.

The different effects of the presence of additive outliers and inliers are analyzed based on Monte Carlo simulations. We analyze three different data generating processes. First, we generate additive outliers in a discrete way as in Vogelsang (1999) and Perron and Rodríguez (2000). Second, additive outliers are generated according to a Bernoulli distribution as was proposed by Franses and Haldrup (1994) and Vogelsang (1999). Third, we generate additive inliers in similar fashion to Cati et al. (1999). The results depend on the nature of each data generating process. Overall, we observe that additive outliers and inliers affect the bias and the MSE of the estimated frcational parameter. We also find size distortions for the t-statistic of the null hypotheses that $d = 0$ or $d = 1$. The size of the additive outliers, the probability of occurance of additive outliers and a drift parameter in data generating process have important effects on the estimated fractional parameter, depending on the value of the true fractional parameter.

Contents

1 Introduction

The analysis of the persistence of some time series has been a topic of sustained research, specially in the macroeconomic context. The paradigm of trend stationarity versus difference stationarity, which involves thinking of persistence as transitory effects versus permanent effects,has been one way to look at this topic. In the middle of this paradigm exists the analysis of persistence using the fractional integration framework. In this context, the integration parameter can be a fraction (not only zero or unity) and then, the effects of some shocks can have a long duration on the behavior of some time series.

Many applications of this approach have been proposed in the literature. Recently,many methods have also been proposed to estimate the fractional parameter. There exist methods in the time domain and in the frequency domain. Another possible area of study is to consider methods based on ordinary least squares (OLS) or maximum likelihood estimation (MLE). In this paper, we are only interested in one particular method related to the frequency domain based on an OLS estimation.

Very often, time series are contaminated due to the presence of aberrant observations which do not necessarily originate in the true data generating process but in time series modeling. These observations are called outliers. There are many kinds of outliers but in this case we are only interested in additive outliers, which by definition affect the level of time series but not the disturbance term associated with it. The effects of additive outliers on the behavior of the unit root tests and $ARMA$ models have been studied before (see for example, Franses and Haldrup 1994, Vogelsang 1999, Perron and Rodríguez 2000, Peña 1990, Chen and Liu 1993, among others).

Another kind of aberrant observation is called an inlier. In this case, the aberrant observation is more likely to be generated by the data generating process. The kind of observation and its impact on behavior of the unit root tests and some measures of persistence have been studied by Cati et al. (1999). The example taken by these authors has been the Brazilian inflation, which is a time series contaminated by many aberrant observations. They consider these observation to be additive inliers associated directly with many government shock programs applied in Brazil to stop high inflation. Whenever inflation was very high, an intervention by the government brought inflation down for a short period of time and the peak can be considered as an additive inlier.

In this paper, we are interested in analyzing the effects on the estimation of the fractional parameter, d, of the existence of additive outliers or additive inliers. As noted above, there are many methods of estimating the fractional parameter, but here we are only interested in analyzing the method of estimation proposed by Geweke and Porter-Hudak (1983), which is a method based on the frequency domain. The different effects of the presence of additive outliers and inliers are analyzed based on simulation evidence. We analyze three different data generating processes. First, we generate additive outliers in a discrete way as in Vogelsang (1999) and Perron and Rodríguez (2000). Second, additive outliers are generated according to a Bernoulli distribution as was proposed by Franses and Haldrup (1994) and Vogelsang (1999). Third, we generate additive inliers in same fashion as in Cati et al. (1999). The results depend on the nature of each data generating process. Overall, we observe that additive outliers and inliers affect the bias and the MSE of the estimated fractional parameter. We also find size distortions for the t-statistic of the null hypothesis that $d = 0$ or $d = 1$. The size of the additive outliers, the probability of occurance of additive outliers and the drift parameter in the data generating process are found to have have important effects on the estimated fractional parameter, depending on true value of the fractional parameter.

The plan of this paper is as follows. Section 2 provides a brief review of the origin of long memory models discusses the properties of long memory models and $ARFIMA$ processes,presents the estimation method of Geweke and Porter-Hudak, and reviews some of the more important of the model of fractional integration in economics anf finance. Section 3 discusses additive outliers and inliers. Section 4 presents the Monte Carlo design for different data generating processes and a description of the results. Section 5 provides the concluding remarks.

2 Long Memory Models

2.1 Origin

The idea of long memory models appear to have developed its roots in the physical sciences, specifically in hydrology early in the 1950s. A particular reference in this respect is Hurst (1951). The study was intended to solve the problem of determining the storage capacity required in the Great Lakes of the Nile River Basin to ensure that the reservoir would be able to supply an irrigation system for the agricultural land in Egypt and Sudan with water throughout the year. Data were collected for the time span of 100 years as the basis of the study's estimates. This most single reference of Hurst (1951) is ascribed to the effect of very long term autocorrelation in observed time series analysis, which later came to be known as a long memory process.

This astounding model of a long memory process has, indeed, attracted the attention of econometricians since around 1980. The pioneering work in this area of econometrics includes Taqqu (1975), Granger and Joyeux (1980), Granger (1981), Hosking (1981). They laid the foundations of the population characteristics of the now widely used Autoregressive Fractionally Integrated Moving Average ($ARFIMA$) process. Geweke and Porter-Hudak (1983) is another basic work on the problem of estimation of the long memory parameter in which a considerable interest is found later in econometric research. Noting this fact, Baillie and King (1996) remark that: "in recent years research activity has gathered momentum with important contributions being made on estimation and testing of a wide range of long memory processes, together with many interesting and imaginative applications over a variety of different fields of economics and finance."[pp. 1]

2.2 Definitions

Long memory is defined as a series having a slowly declining correlogram or equivalently an infinite spectrum at zero frequency. Denoting ω a particular frequency, Geweke and Porter-Hudak (1983) quoted the following definition of long memory models given by McLeod and Hipel (1978):

> *"attention has recently been given to two single-parameter models on which the spectral density function is proortional to ω^{-r}, for $1 < r < 2$ for near to the frequency 0 and the asymptotic decay of the autocorrelation function is proportional to τ^{r-1}. because the spectral density is unbounded at $\omega = 0$, equivalently, the autocorrelation function is not summable—these are long memory models."[pp. 221]*

We can follow Granger and Ding (1996) to explain the definition and the properties of long memory models. Let h_j, $j = 0, 1....$,be an infinite sequence of constants and define a time changing sequence of filters that is

$$h_t(L) = \sum_{j=0}^{t} h_j L^j, \tag{1}$$

where L is the backward operator. A time series $\{x_t\}$ can be generated by

$$x_t = h_t(L)e_t, \tag{2}$$

where the input series is supposed to have the form

$$e_t = m + \varepsilon_{t,} \tag{3}$$

where ε_t is white noise sequence with mean zero and variance σ_ε^2. Denoting $m(t) = mh_t(1)$, we can define:

$$v(t) = \sigma_\varepsilon^2 \sum_{j=0}^{t} h_j^2 \tag{4}$$

$$s_t = \sum_{j=0}^{t} h_j \varepsilon_{t-j}, \tag{5}$$

$$f_T(\omega) = E[T^{-1} \mid \sum_{t=1}^{T} s_t e^{it\omega} \mid^2]. \tag{6}$$

If $f(\omega) = \lim_{T\to\infty} f_T(\omega)$ exists for all $\omega \geqslant 0$, then $f(\omega)$ is called the spectrum of x_t. If we confine our attention to $m(t) = mh_t(1)$, we can say that $m(t)$ will be the deterministic trend in mean and $v(t)$ will be the variance of x_t, which holds provided $h_t(1)$ is monotonically increasing. It is also evident here that these two terms ($m(t)$ and $v(t)$) may tend to constants. These properties are discussed in Granger (1988) and Granger and Ding (1996).

For large T, $f_T(\omega)$ serves as a close approximation to $f(\omega)$ for almost all frequencies, so that the series x_t is called 'long memory',which holds provided the spectrum $f(\omega)$ tends to infinity as $\omega \to 0$, and $f(\omega)$ is bounded above for all except a finite number of other values of ω. These properties are discussed by Granger and Joyeux (1980) and Hosking (1981). In the case of the spectrum which tends to infinity, the sequence h_j needs to be a divergent one. For $h_j = A_j^{d-1}$, the series generates the $I(d)$ process. This process possesses the spectrum which is proportional to ω^{-2d} for small ω . In this case,we have a fractionally integrated process provided d is not an integer. It has a 'long memory' property because a shock ε_t at time t

will continue to influence future values of x_t (i.e., x_{t+k}) for a larger k than would be the case for the standard stationary $ARMA$ process, for which h_j declines exponentially.

2.3 The ARFIMA Processes

Kennedy (1998) presents a technical note on ARFIMA processes which states that the concept of fractional integration is useful in testing for non-stationarity. There are two kinds of models to examine in this respect. One is $(1-\alpha L)y_t = \varepsilon_t$ where $\alpha \geqslant 1$ and the other is $(1-L)^d y_t = \varepsilon_t$ where $d \geqslant 0.5$, both of which correspond nonstationarity. In the second case, d takes on non-integer or fractional values, from which follows the term 'fractional'. In order to understand the rationale for the term 'integrating', consider a random walk model such as $y_t = y_{t-1} + \varepsilon_t$, which after some algebra can be written as $y_t = y_0 + \sum \varepsilon_{t-i}$. It is clear here that the variable y is determined by summing up or integrating the error terms. In this way, $ARIMA$ turns into fractional $ARIMA$ i.e. an $ARFIMA$ process.

Sowell (1992) and Crato and Rothman (1994) provided explanations for why an $ARFIMA$ process is preferable to trend stationary and difference stationary series. Kennedy (1998) also provides a brief discussion of the advantages of $ARFIMA$ processes over the unit root based tests for non-stationarity. He stated:

> *Although there are higher computational costs, modeling in terms of fractional integration has several advantages. It allows a continuous transition from non-unit root behavior to a unit-root, it is better suited to capturing low frequency (long memory) behavior and so is able more adequately to model on-term persistence and it nests both difference stationary and trend stationary models. In short, it provides a more flexible alternative against which to test unit roots and because of this empirical studies using this*

approach tend to reject a unit root.[pp. 286]

The $ARFIMA$ process, however, is not free of limitations. In the case of parametric expressions, the standard $ARFIMA(p,d,q)$ process for the infinite autoregressive (AR) and moving average (MA) representation is relatively complicated. Chung (1996) provides some alternative methods for the calculation of these autocorrelations. In this context, Baillie (1996) remarks that estimators produced by the $ARFIMA$ process are conceptually attractive, but the performance of these estimators is not good when one compares the theoretical aspects with simulation work and it produces the same result irrespective of the nature of the white noise such as low-order or high-frequency dynamics. Another drawback to this kind of method is that it is difficult to give an intuitive interpretation to a non-integer difference. A major problem, in practice, consists of distinguishing between a long memory process and a non-stationary process. For these reasons, models of fractional integration have not been widely used and are not included in any standard econometric software packages.

In case of long term autocorrelation, several models show persistence in shocks and provide an extension of the concept of nonstationarity. The most common model of this kind is known as the fractionally integrated white noise series which is as follows:

$$(1-L)^d y_t = \varepsilon_t$$

where $y_t = Y_t - \mu$. This time series has an infinite moving average representation if $\mid d \mid < 0.5$, but it is nonstationary for other values. An extension of this model is

$$(1-L)^d[y_t - \gamma_1 y_{t-1} - \dots\dots\dots\dots - \gamma_p y_{t-p}] = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots\dots - \theta_q \epsilon_{t-q}$$

which is termed the $ARFIMA$ (p,d,q) model.

2.4 The Method of Geweke and Porter-Hudak

There are a variety of ways of estimating the parameter d. The statistical measure of long memory used by Hurst (1951) and later used by Mandelbrot (1972, 1975) was the rescaled range statistic which has, according to Granger and Joyeux (1980), little intuitive appeal and thus is beyond the scope of our study. Here,we are essentially interested in the estimation procedures proposed by Geweke and Porter-Hudak (GPH) (1983), which generalized the definitions of fractional Gaussian noise and integrated or fractionally differenced series and showed that the two concepts are equivalent. This procedure is based on an estimation in the frequency domain. For the model $(1-L)^d x_t = \varepsilon_t$,where$\{x_t\}$ is assumed to be a time series process, $d \in (-.5, .5)$ and ε_t is serially uncorrelated, the spectral density of the time series $\{x_t\}$ is

$$f_2(\omega; d) = (\sigma^2/2\pi)\left|1 - e^{-i\omega}\right|^{-2d} = (\sigma^2/2\pi)\{4\sin^2(\omega/2)\}^{-d}. \qquad (7)$$

A time series with the spectral density $f_2(\omega; d)$ is called an integrated or fractionally differenced series which suggests that $\lim_{\omega\to 0} \omega^{2d} f_2(\omega; d) = \sigma^2/2\pi$ and the autocorrelation function (for $d \neq 0$) is $\rho_2(\tau; d) = \Gamma(10d)\Gamma(\tau + d)/\Gamma(d)\Gamma(\tau + 1 - d)$, which leads to $\lim_{\tau\to\infty} \tau^{1-2d}\rho_2(\tau; d) = \Gamma(1-d)/\Gamma(d)$.

Now consider $(1-L)^d y_t = u_t$, where u_t is a linear and stationary distributed process with the spectral density function $f_u(\lambda)$ which is supposed to be finite, bounded away from zero and continuous on the interval $[-\pi, \pi]$. Based on GPH, one has

$$log\{f_y(\omega_j)\} = \log\{f_u(0)\} - d\log\{4\sin^2(\omega_j/2)\} + \log[f_u(\omega_{j)}/f_u(o)]. \qquad (8)$$

According to GPH we can estimate d from a regression based on (8) using spectral ordinates $\omega_1, \omega_2.......\omega_m$, from the periodogram of y_t, that is $I_y(\omega_j)$:

$$log\{I_y(\omega_j)\} = a - d\log\{4\sin^2(\omega_j/2)\} + v_j, j = 1.....n \qquad (9)$$

where

$$v_j = \log[f_u(\omega_j)/f_u(0)] \tag{10}$$

and v_j is supposed to be *i.i.d.* with zero mean and variance $\pi^2/6$. Thus the least square estimator of d is asymptotically normal. If the number of ordinates n is chosen such that $n = g(T)$, where $g(T)$ is such that $\lim_{T\to\infty} g(T) = \infty$, $\lim_{T\to\infty}\{g(T)/T\} = 0$ and $\lim_{T\to\infty}\{(\log(T)^2/g(T)\} = 0$, then the OLS estimator of d in (9) takes the limiting distribution as follows:

$$(\widehat{d} - d)/\{var(\widehat{d})\}^{1/2} \Rightarrow N(0,1) \tag{11}$$

When the OLS estimator d is significantly different from zero, the sample of the specific size is fractionally integrated. Here, in our estimation, we use $n = g(T) = \sqrt{T}$.

2.5 Applications of Long Memory Models

Long memory models originated in the domain of the geophysical sciences. One interesting application in this domain is in the detection of the possibility of climatic change. Baillie (1996) mentions that there have been several findings which showed an apparent upward trend in world temperatures since the mid of the nineteenth century (see e.g. Seater 1993). There are also many applications related to macroeconomics. For example, there is a long standing debate in macroeconomics as to whether real GNP is difference stationary or trend stationary. Long memory models have an important role to play in resolving this kind of dilemma. (see for example Diebold and Rudebusch 1989 and Sowell 1992).

The long memory property is also present in price series over a long span of time. The Wheat Price Index of Beveridge (1925) is an example in that respect. The model also has an extremely important application to the volatility of asset prices (e.g Ding, Granger and Engle 1993) and in stock market returns (e.g. Lo 1991). These models apply the rescaled range statistic to test for long memory in stock market returns. It is also applicable to exchange rates (Cheung 1993). Finally,it can be used to model the premium and to analyze the behavior of a set of interest rates. Shea (1991) applies the GPH procedure in this context.

3 Additive Outliers and Inliers

Since Fox (1972) introduced the notions of additive and innovational outliers, issues related to these type of atypical observations in time series have received considerable attention in the statistics and econometrics literature. The outlier detection issue, itself, has received particular attention. Another topic of interest in the research has been the estimation of $ARMA$ models in the presence of outliers. In this case, as mentioned by Chen and Liu (1993), a common approach is to identify the location and the types of outliers and then to accommodate the effects of outliers using intervention models as proposed by Box and Tiao (1975). This approach requires iterations between the stages of outlier detection and estimation of the model.

An outlier is a realization of the time series that is an aberration, that is, not from the usual data generating process. In the literature (see Fox 1972, Box and Tiao 1975) two kinds of outliers have been discussed. The first class is the additive outlier, where the aberrant observation affects directly the level of the time series. The second class is the innovative outlier, where the aberrant observation affects the noise component of the time series. In other words, the effects on the time series of these kinds of observations are immediate or gradual, respectively.

In the context of integrated data (a process with an autoregressive unit root), the effects of additive outliers have recently been the object of sustained research. It is by now well recognized that outliers affect the properties of unit root tests (e.g., Franses and Haldrup 1994). They do so by inducing a negative moving average component in the noise function, which causes most unit root tests to exhibit substantial size distortions towards rejecting the null hypothesis too often. By " size distortions" we are referring to the difference between the nominal and exact size of a particular test. As is well known, the size of a test is defined as the probability of rejecting the null hypothesis when it is true. Hence, the nominal size is the probability of rejecting the null hypothesis implied by the pre-specified level of significance upon which the critical values used to test hypothesis are based. In most of cases, this level corresponds to 5.0%. The exact size of a test is the actual probability of rejecting the null hypothesis, which is usually measured artificially generated data. Consider the following example. Suppose that using artificial data and critical values at 5.0%, the percentage of times that a statistic rejects the null hypothesis when it is true is found to be 12.0%. This value corresponds to the exact size and the nominal size is 5.0%. Evidently, this statistic displays size distortions because exact size is higher than the nominal size. In this case, we say that this statistic is too liberal. If the exact size is less than the nominal size, we say that the statistic is too conservative.

The introduction of a strong negative moving average components implies frequent rejectio of the null hypothesis of a unit root. This is because although the time series has a unit root, the presence of negative moving average errors allows it to behave in a manner similar to a stationary time series. The unit root coefficient is biased to zero and the t-statistic will reject the null hypothesis of a unit root. Normally, the solution for this problem is to use a longer lag in the application of the ADF statistic. However this is not always possible because it depends of the size of the additive outliers. Higher additive outliers imply moving average components closer to the unit circle (in this case, closer to -1) and then the process is not invertible and using a very large lag is not useful. The suggestion of Franses and Haldrup (1994) was to apply Dickey-Fuller (1979) unit root test by incorporating dummy variables in the autoregression chosen on the basis of the outlier detection procedure proposed by Chen and Liu (1993).

In an interesting recent paper, Vogelsang (1999) makes two contributions to the issue of the effects of additive outliers on unit root tests. First, recognizing that outliers induce a negative moving average component, he suggests using unit root tests developed by Stock (1990) and Perron and Ng (1996) that are robust, in terms of achieving exact size close to nominal size in small samples, even in the presence of a substantial negative moving average component. He shows, via simulations, that these unit root tests are little affected by systematic outliers. Second, he recognized that one can take advantage of the null hypothesis of a unit root in devising an outlier detection procedure. This allows the derivation of a non-degenerate distribution for the t-statistic on the relevant one-time dummy variable.

In similar context and related to the paper of Vogelsang (1999), Perron and Rodríguez (2000) have contributed to the outlier detection issue suggesting two modifications to the original method proposed by Vogelsang (1999). In the first method, they suggest using modified asymptotic critical values in the iterative outlier detection procedure, which lead to better size properties for the test procedure. In the second contribution, they suggest using the first differences of the data. This last alternative has been shown to result in better size and power properties for the test to detect for additive outliers.

The detection procedure, suggested by Vogelsang (1999), starts with the following regression estimated by OLS (if necessary, a time trend can also be included),

$$y_t = \widehat{\mu} + \widehat{\delta} D(T_{ao})_t + \widehat{u}_t \tag{12}$$

where $D(T_{ao})_t = 1$ if $t = T_{ao}$ and 0 otherwise. Let $t_{\widehat{\delta}}(T_{ao})$ denote the t-statistic for testing $\delta = 0$ in (12). Following Chen and Liu (1993), the presence of an additive outlier can be tested using

$$\tau = \sup_{T_{ao}} \mid t_{\widehat{\delta}}(T_{ao}) \mid .$$

Assuming that $\lambda = T_{ao}/T$ remains fixed as T grows, Vogelsang (1999) showed that as $T \to \infty$,

$$t_{\widehat{\delta}}(T_{ao}) \Rightarrow H(\lambda) = \frac{W^*(\lambda)}{(\int_0^1 W^*(r)^2 dr)^{1/2}} \tag{13}$$

where $W^*(\lambda)$ denotes a demeaned standard Wiener process (i.e. $W^*(\lambda) = W(\lambda) - \int_0^1 W(s)ds$). If (12) also includes a time trend, $W^*(\lambda)$ will denote a detrended Wiener process. Furthermore, from the continuous mapping theorem it follows that,

$$\tau \Rightarrow \sup_{\lambda \in (0,1)} |H(\lambda)| \equiv H^*. \tag{14}$$

The distribution given in (14) is non-standard but is invariant with respect to any nuisance parameters, including the correlation structure of the noise function. The asymptotic critical values for τ were obtained using simulations (see Vogelsang 1999, Perron and Rodríguez 2000).

The outlier detection procedure recommended by Vogelsang (1999) is implemented as follows.[1] First, compute the τ statistic for the entire series and compare τ to the appropriate critical value. If τ exceeds the critical value, then an outlier is detected at date $\widehat{T}_{ao} = \arg\max_{T_{ao}} |t_{\widehat{\delta}}(T_{ao})|$. The outlier and the corresponding row of the regression is dropped and (12) is again estimated and tested for the presence of another outlier. This continues until the test shows a non-rejection.

Notice, however, that this paper does not apply procedures to detect for additive outliers. We are interested in a earlier step, that is, in understanding the effects of the additive outliers on some specific parameters.

[1] It is equivalent to the *stepwise* procedure to select for multiple outliers. See Hawkins (1980).

Fractional integrated processes can be considered to be a bridge between stationary and non-stationary cases. This framework offers a wider spectrum of possibilities regarding the possible behavior of a time series. There are many time series for which the application of traditional unit root tests does not allow us to arrive a clear conclusion about the properties of these time series. When a statistic where null hypothesis is stationarity is applied, contradictory conclusions can be found. The reason can be the fact that the time series effectively follows a fractionally integrated process and hence, cannot be modeled as either $I(0)$ or $I(1)$. Because fractional integration (I(d)) is a bridge between I(0) and I(1) processes, the presence of additive outliers can affect the behavior of the estimates in this framework as a consequence of the introduction of moving average components. In this context, we would like to know the effects of the presence of additive outliers on the estimation of the fractional parameter. With respect to unit root tests, for example, we know that the estimated autoregressive coefficient is biased towards zero when there are additive outliers and hence strong size distortions of these tests are found. With respect to fractional integration similar effects are likely but more detailed evidence about the behavior of the estimates is reasured because we have many possible values of the fractional parameter between zero (stationary case) and unity (non-stationary case). Also, it is know that for fractional parameters less than 0.5, the processes are stationary and for other cases (more than 0.5 but less than unity) the processes are non-stationary. Then, another question is to answer in which case the presence of additive outliers introduces more distortions in the estimates of the fractional parameter.

Our intuition here is simply a generalization of what does happen in the unit root case. When the fractional parameter is closer to unity (non-stationary case), we can expect more distortions. Conversely, when the fractional parameter is closer to zero (stationary case), fewer distortions are likely to prevail. Similar behavior is possible when the fractional parameter is greater than -0.5 but less than zero.

The presence of inliers is another factor that can affect estimation of the fractional parameter. Inliers can also be considered as aberrant or atypical observations (like additive outliers) but generated from the data generating process (unlike additive outliers). The issue of inliers was treated by Cati et al. (1999) in the context of unit root tests with an empirical application to the Brazilian inflation. Cati et al. (1999) consider inliers as equivalent to the observations associated with the application of government shocks in order to stop the high inflation process, as was the case in the Brazilian economy. Essentially government shocks can be seen as creating inliers whose magnitude is related to the current level of the series. As is the case with unit root tests, we suppose that the presence of these inliers will contaminate the estimates of the fractional parameter by reducing its magnitude and then showing a tendency towards the stationary case. This argument is essentially the flip-side of the argument of Perron (1989), who argued that permanent changes in the trend function of a series with a stationary noise biases standard unit root tests towards accepting the unit root hypothesis and concluding that shocks have persistent effects. Here, in the case of inliers, temporary, but large, changes bias these measures in the opposite direction.

The idea behind the relationship between inliers and the unit root hypothesis is the fact that estimates of the autoregressive parameter will be biased towards zero and then size distortions will be found when we use standard unit root tests. A similar hypothesis can be suggested for the case where the fractional parameter is estimated but more details about the behavior of this parameter (stationary or non-stationary case) are needed.

Additive outliers and inliers are frequently observed in time series data. To mention only some examples, there is the time series of inflation in Brazil analyzed by Cati et al. (1999). Inflation in other Latin American countries (such as Argentina, Bolivia, Chile and Peru) are other examples analyzed by

Rodríguez (2000). Vogelsang (1999) and Franses and Haldrup (1994) used as examples the evolution of the real exchange rate between Finland and U.S. to show that presence of additive outliers is relevant.

4 The Monte Carlo Design

A theoretical treatment of this topic is not our goal in this paper. Instead, a standard procedure for dealing with this type of topic is to analyze the results obtained from a simulated experiment. In other words, this means the use of a Monte Carlo experiment. A Monte Carlo experiment consists of generating repeated samples of artificial data (according to some characteristic or properties) for some sample size, and then analyzing thebehavior of the relevant statistics. In our case, for example, we are interested in analyzing the behavior of the estimates of the fractional parameter. One way to do this is to calculate some characteristics of this estimate such as the Mean Square d Error (MSE) and the bias. When the size and power of one statistic is the principal focus, we calculate the number of rejections of the null hypothesis found in all the replications used.

Three data generating process will be considered in the following simulation experiments. We think that these three data generating process are sufficient general to show the principal effects of aberrant observations on the estimates or statistics. The first two data generating processes correspond to the case where additive outliers are considered. The third data generating process considers the case where inliers are present.

The first data generating process is same as that considered in Vogelsang (1999) and Perron and Rodríguez (2000). It involves the case where additive outliers are fixed. It can be defined according to:

$$y_t = n_t + \sum_{i=1}^{m} \delta_i D(T_{ao,i})_t + u_t \qquad (15)$$

$$(1-L)^d u_t = v_t \qquad (16)$$

where $v_t \sim i.i.d.\ N(0,1)$, d is the fractional parameter, $D(T_{ao,i})_t = 1$ if $t = T_{ao,i}$ and 0 otherwise and δ_i is the size of the additive outlier. Four sizes of additive outliers are considered (that is, $m = 4$ in expression (15)), and two different assumptions about their values are considered. In the first case, $\delta_i = 0$ (for $i = 1, 2, 3, 4$). This case will illustrate the situation where we have no outliers, and thus no size distortions and no bias in the estimates will be observed. In the second case, $\delta_1 = 10, \delta_2 = 5, \delta_3 = 2, \delta_4 = 2$. This case will illustrate the effects when there are additive outliers of "large" size. The second specification is close to that used by Perron and Rodríguez (2000). Our goal is to see as clearly as possible the effects of large additive outliers in comparison to a situation where there are no additive outliers. The variable n_t represents the deterministic components. In the experiment, we consider only the case where a constant is included in the regressions; that is, $n_t = \mu$. In the simulations of the expression (15), without loss of generality, we consider the case where $\mu = 0$.

The second data generating process was used in Vogelsang (1999) and Franses and Haldrup (1994) where systematic additive outliers are considered. This process can be expressed by

$$y_t = n_t + \delta_t D(T_{ao})_t + u_t \qquad (17)$$

$$(1-L)^d u_t = v_t \qquad (18)$$

where $D(T_{ao,i})_t = 1$ if $t = T_{ao,i}$ and 0 otherwise and δ_t is generated according to $i.i.d.$ Bernoulli random variables with $\text{Prob}(\delta_t = 1) = 0.5p$, $\text{Prob}(\delta_t = -1) = 0.5p$ and $\text{Prob}(\delta_t = 0) = 1 - p$, with $p \in [0,1]$. P denotes the probability of an outlier occurring at any given time. As p approaches unity, the greater the chance an outlier is realized at a given time. We have chosen values for $p \in [0.3, 0.9]$ with $step = 0.1$, which gives us a total of seven values for p. We think that this set of values for p will provide considerable detail about the effects of this parameter on the behavior of our estimates or size of tests. Finally, similar conditions are specified for v_t and n_t as in the first data generating process.

The third data generating process considers the case where there are shock plans (or inliers) that are short-lived but important in magnitude. In this case, we follow a data generating process similar to that used by Cati et al. (1999), modified to the context of fractional integrated processes. Then, we consider a fractional integrated process interrupted by occasional inliers or shock plans:

$$y_t = \mu + u_t, \qquad for \quad t \notin \{t_{i,j}\} \tag{19}$$

$$(1-L)^d u_t = v_t \tag{20}$$

$$y_t = a, \qquad for \quad t \in \{t_{i,j}\} \tag{21}$$

where $j = 1, 2, ..., s$; $i = 1, 2, ..., n_j$. Here, $t_{i,j}$ refers to the time index of the ith observation of plan j. There are s shock plans and each contains n_j observations. The series is a fractional integrated process with a drift μ except when the plan is in effect, in which case the level of the series drops to a value a. As for other two data generating processes, $v_t \sim i.i.d.\ N(0,1)$. For this data generating process, $\mu \in [-0.8, 0.8]$ with a step of 0.4. The parameter $a = 0.4$. This parameter can be interpreted as the threshold level to which the time series is driven when an additive inlier occurs,or, according to the example of Cati et al. (1999), the level to which the time series of inflation is driven after the application of a government shock. Finally, the integration parameter $d \in [-0.96, 0.96]$ with a $step = 0.24$. We do not include $d = 1$ because this case corresponds to the case of a unit root and has been previously analyzed by Franses and Haldrup (1994) and Vogelsang (1999). The case where $d = -1$ is also not included because it corresponds to the case of an overdifferenced time series.

For all data generating processes, the sample sizes considered are $T = 50, 100,$
200, and 500. We think that these sample sizes are fairly common as in empirical work. For example sample sizes like $T = 50$ or $T = 100$ are

frequently found in macroeconomic research. The other sample sizes are frequently found in finance research. The number of replications considered for each set of parameters is 1000 and a $seed = 12345$ was used. The number of simulations used is similar to those used in the literature. However, another reason for choosing this number was the number of data generating processes and values of the different parameters considered in the design. All simulations were performed in Gauss and our programs were based on programs constructed by Pedro de Lima.[2]

In the following section we will discuss the results about the effects of aberrant observations using three indicators. We are interested in analyzing the bias and the MSE of the estimate of the fractional parameter. The bias is defined as the difference between the true parameter value and the expected value of the estimate. In a formal expression: $bias = E[\widehat{d} - d]$, where E denotes the expectation operator. The MSE is calculated as: $MSE(\widehat{d}) = bias^2 - var(\widehat{d})^2$, where var denotes variance. Finally, we like to know the effects on the calculated size of the t-statistic of the null hypotheses that $d = 0$ and $d = 1$, respectively. To calculate this statistic we use standard formulae. Notice that when the true fractional parameter is different from zero (for example), that is $d = 0$, and we calculate the t-statistic for the null hypothesis that $d = 0$ (i.e. $t_{d=0}$), we are analyzing the power of the test. Hence, overall we will use the term size but notice that it includes in some cases the notion of power. In all cases, we are analyzing two-tailed tests.

Finally we need to mention several points before presenting the description of the results. In following analysis, frequently we mention that the test has the "correct" size. By that we mean that the exact size is very close or equal to the nominal size of 5%. Also, we have concentrated our analysis on the method proposed by GPH (1983). This is because is the most widely used method of estimating the fractional parameter. It is the morst intuitive

[2]The Gauss Program of Pedro De Lima appears in the Webpage of the Department of Economics of John Hopkins University, USA.

method because it is based on an OLS regression. Of course there are many other methods proposed in the literature and a possible comparison between them with respect to the effects of aberrant observations can be considered a project for further research.

4.1 Description of the Results

Firstly, we consider results for the case where outliers are generated in a simple additive manner; that is, the first data generating process in the last section. Tables 1a-1d present the results for the case where there are no additive outliers. In terms of the bias and the MSE, there are no significant variations for all of the fractional parameter values considered. We only observe that the MSE is different for the two extreme values of d. In fact when d is close to unity the MSE is smaller. This is because the bias and the variance are smaller, probably as a consequence of a better estimation of the fractional parameter in opposition to the case where d is close to -1.

When the true parameter $d = 0$, the exact size is closer to the nominal size when sample size increases, which is the result expected. When the true fractional coefficient is closer to -1, we strongly reject the null hypothesis that the fractional coefficient is equal to zero. On the other hand, as we can expect, when the true fractional coefficient is close to unity, it is very difficult to reject the null hypothesis that the coefficient is different from one. This is also true for very large sample sizes such as $T = 500$. It is consistent with results found for the case of unit root tests, which have no power even for higher sample sizes.

Tables 2a-2d present the results for the case when there are some large and small additive outliers. For the samples sizes of $T = 50, 100$ and 200, the exact size for the null hypothesis that $d = 0$ is closest to zero when the true fractional parameter is closer to -1. The reverse is true when the true fractional parameter is closer to unity. The opposite arises for the exact size of the null hypothesis that the fractional parameter is equal to unity.

However, we observe that there is almost no size distortion when we use $T = 500$, a sample size, unfortunately, that is not frequently available in macroeconomic applications.

The results with respect to the bias and the MSE are also related to the behavior of the true fractional parameter. In fact, when this parameter is closer to -1, bias and MSE increase. The reverse is true when we have $d > 0$ but less than unity. For this sample size, the bias is important when $d < 0$. Although bias and MSE are smaller for $T = 500$, the exact size of the t-statistic of the null hypothesis that $d = 1$ is higher compared to other sample sizes. Given that in this case d is close to unity, this behavior is similar to the power problems observed for most of unit root tests in the literature.

Tables 3a-3c present the results when the additive outliers have been generated according to a Bernoulli distribution; that is, following the second data generating process presented in the last section. When the fractional parameter is very close to -1 (for example -0.96), the MSE is higher for intermediate values of the probability parameter p. Also, the bias of the estimator is higher when the probability parameter is smaller. With respect to the size of the t-statistic for the null hypothesis that $d = 0$, the exact size is more distorted when the probability parameter is closer to unity and the fractional parameter is closer to -1. When the fractional parameter is closer to zero, we observe smaller size distortions. When the fractional parameter is closer to unity, the exact size of the t-statistic for $d = 0$ is more distorted when the probability parameter is closer to unity. For a particular value of the probability parameter, size distortions are higher for cases where $d < 0$. For $d = 0$ or $d > 0$, the bias is negative which means that there is an underestimation of the true value of d. The bias and the MSE are very stable for all values of d and both quantities are smaller when $d < 0$.

When sample size increases to $T = 100$ we can observe that the bias and the MSE increase when d goes to -1, specially when p is smaller. At the same time, the exact size of the t-statistic for $d = 0$ is smaller when d goes to -1. For $d > 0$, the bias and the MSE are smaller than in the case where $T = 50$. Until $d = 0.72$, the size distortion (non- rejection) of the t-statistic for $d = 1$ is higher. For $d = 0.96$, the exact size of the t-statistic for $d = 0$ is correct bu this is not the case for the t-statistic for $d = 1$, which is distorted towards non- rejection.

When $T = 200$ and $d < 0$, the rejection of the null hypothesis that $d = 1$ appears to be correct. The bias and the MSE decrease faster when we pass from $d = -0.72$ to $d = -0.48$. For $d > 0$, the exact size of the t-statistic for $d = 1$ appears to be correct and we observe a problem with the size of the t-statistic for $d = 0$. This problem is less severe when $d = 0.48$, for example. But from $d = 0.72$, the problem is that it is very difficult to reject the null hypothesis that $d = 1$. For $d = 0.96$, we observe a high rate of rejection of the null hypothesis that $d = 0$, and it is very difficult to reject the null hypothesis that $d = 1$.

Four conclusions can be derived from this set of tables. First, the bias and the MSE are not severely affected when the probability parameter increases if sample size increases at the same time. Second, the exact size of the null hypothesis that $d = 1$ is not severely affected when the probability parameter increases or decreases. The difficulty in rejecting this null hypothesis is related to the simple fact that the fractional parameter is too close to the unit circle. But the situation is different situation for the t-statistic that $d = 0$, especially when $d < 0$. Third, size distortions do not decrease monotonically when sample size increases. Fourth, the effect of the parameter p depends on or is related to how close the fractional parameter is to the unit circle.

Results for the case where there exist additive inliers are presented in tables 4a-4c. Consider, for example, the case where $T = 100$ and $d = -0.96$. For this case, we observe that the bias and the MSE are important but their magnitude does not change for different values of μ or a. That is, there is no particular effect of the dimension of the drift parameter or the threshold parameter on the behavior of the bias and the MSE for a particular size of the fractional parameter and a particular sample size. In this case, the size distortions are also not very strongly affected, especially for the case of the t-statistic for the null hypothesis that $d = 1$.

Yet, we observe that the bias and the MSE decrease when the fractional parameter is closer to zero. When $d = 0$, again, we do not observe particular effects on the bias and the MSE when the values of μ or a are changing. The same conclusion is obtained for the size distortions. When the fractional parameter $d > 0$, the bias and the MSE are higher. In particular, the bias is more negative in value, which means that the true fractional parameter is underestimated. In this case, we observe some changes in the behavior of the bias and the MSE when the drift parameter goes to unity or to minus unity. In these cases, the bias and the MSE are higher. The problem is not observed when $a = 0$ or $a = 4$, which means that the threshold parameter has no important effects on the estimation of the fractional parameter.

Similar behavior is observed in the size distortions of the t-statistics. In effect, there are greater distortions when the drift parameter goes to unity or to minus unity. The size distortion for the t-statistic of the null hypothesis that $d = 1$ is more evident when the drift parameter is zero and when $a = 0$ or $a = 4$. When the fractional parameter is closer to unity, there is no possibility of rejecting the null hypothesis that $d = 0$, especially when the drift parameter goes to unity or to minus unity. A similar problem in rejecting the null hypothesis that $d = 1$ is observed, especially when the drift parameter is equal to zero and $a = 0$ or $a = 4$.

A different result in observed when the drift parameter is different from zero. In fact, when the drift parameter goes to unity or to minus unity, the exact size of the t-statistic of the null hypothesis that $d = 1$ is correct, including the case when the fractional parameter is closer to unity.

Now, consider the results for the case where $T = 200$. The bias and the

MSE decrease when the fractional parameter is closer to zero which is due to the increase in the sample size (in comparison with the results that we observed for $T = 50$ or $T = 100$). The bias and the MSE do not suffer from severe changes when the drift parameter or the parameter a are higher. We can observe some small problems in rejecting the null hypothesis that $d = 0$. When $d = 0$, there are no real problems with the bias and the MSE of the fractional parameter. There is also no major problem with the exact size of the t-statistics.

When $d < 0$, the bias and the MSE seem to behave in a stable fashion but the bias is more in negative value for all values of μ and a. When $d > 0$, the exact size of the t-statistic for $d = 0$ is close to unity when the true fractional parameter is closer to unity. On the other hand, the exact size of the t-statistic for $d = 1$ decreases when the true fractional parameter is closer to unity. We observe many difficulties when the true drift parameter is zero.

The MSE increases when the fractional parameter is closer to unity and for this case, we observe a strong rejection of the null hypothesis that $d = 0$ but not a clear reject of the null hypothesis that $d = 1$.

Finally, some issues that pertain to all sample sizes need to be mentioned here. First, when we compare Table 2 with Table 1, it is possible to observe the clear effects of additive outliers against a case where there do not exist any aberrant observations. The evidence with respect to higher bias and MSE is clear. Also, we can observe that there are size distortions for the t-statistic of the null hypothesis that $d = 0$. Similar conclusions can be extracted from a comparison of Tables 3 or 4 with Table 1.

Second, when the two kinds of additive outliers are compared (Tables 2 and 3), the exact size of the null hypothesis of $d = 0$ or $d = 1$ is very similar for the two cases. However, we observe some differences in the values of the bias and the MSE. In fact, results obtained for the second data generating process show that the magnitude of bias and the MSE are smaller compared to the results from the first data generating process.

Third, when we compare the results from the third data generating process against the first data generating process (Tables 4 and 1), conclusions similar to those mentioned above can be drawn. We can also add the fact that for additive inlier we can observe that the size distortions are marginally higher in comparison to those found in the Table 2.

5 Conclusions

The primary motivation of this paper is to analyze the effects of additive outliers and inliers on the behavior of the estimated value of the fractional parameter. With this end in view, an extensive set of simulations employing two different methods of generating additive outliers were considered. Our principal criteria for analyzing the behavior of the estimated parameter are the bias, the MSE and the exact size of the t-statistic of the estimated fractional parameter.

When there are additive outliers generated according to the first data generating process, the exact size of the t-statistic of the null hypothesis that $d = 0$ goes to zero, especially when the true fractional parameter is negative. The reverse situation occurs for the t-statistic of the null hypothesis that $d = 1$ when the true fractional parameter is positive. In respect of the bias and the MSE, they are more affected for the size of the fractional parameter.

When the additive outliers are generated according to a Bernoulli distribution, the t-statistic for the null hypothesis that $d = 0$ is distorted when p is higher, particularly when $d < 0$. Clearly, the bias and the MSE are affected in an inverse manner, because they increase when p decreases.

Finally, when there exist additive inliers, the bias and the MSE are higher when the drift parameter goes to unity or to -1, specially when $d > 0$. On the one hand, we cannot observe particular effects on the exact size of the t-statistics caused by the behavior of the parameter a, which means that the threshold parameter has no important effects on the estimation of the fractional parameter. Following the example of inflation analyzed by Cati et al. (1999), we can say that the level to which inflation falls after the application of a government program (an inlier) has no bearing on the estimation of the fractional parameter or the persistence of the inflation process. Size distortions for the t-statistic of the null hypothesis that $d = 0$ are observed and they are higher when the drift parameter is close to the unit circle. On the other hand, the exact size of the t-statistic of the null hypothesis that $d = 1$ is more affected when $\mu = 0$.

Comparison between a situation in which there exist additive outliers (discrete or systematic) or inliers to a situation in which these kinds of observations don not exist, showed that they have particular and important effects on the estimation of the fractional parameter and on the estimation of the t-statistic for verifying the null hypothesis that $d = 0$ or $d = 1$. Overall, we observe clear size distortions and the bias and the MSE are higher or clearly different with respect to the situation where these kinds of observations do not exist.

There are small differences between the kinds of distortions observed when we compare the two kinds of additive outliers or also when we compare the results obtained from additive inliers against additive outliers.

Our results suggest that researchers should be careful when they thins that a time series is affected by additive outliers or inliers. Inference about the fractional parameter or hypothesis tests based on the traditional t-statistic will give misleading conclusions if these kinds of observations are not taken into account. Hence, a first step in applied research should be to apply some methods of detecting additive outliers. We know that this step is not present in this paper. But, our goal was to verify whether or not these kinds of aberrant observations posed problems in this particular kind of model. Given that the results show that there are important effects, the next step in future research will be to propose a method of detecting additive outliers in this context. Approaches proposed by Vogelsang (1999) and

Perron and Rodríguez (2000) will be considered. Finally, we plan to extend our analysis to other methods of estimating the fractional parameter.

References

[1] Baillie, R.T. (1996), "Long Memory Processes and Fractional Integration in Econometrics," *Journal of Econometrics* ***73*** *(5-59).*

[2] Baillie, R.T. and Maxwell L. King (1996), "Editors' Introduction: Fractional Differencing and Long Memory Processes," *Journal of Econometrics* **73**, 1-3.

[3] Beran, J.A. and N. Terrin (1994), "Estimation of the Long Memory Parameter, Based on a Multivariate Central limit theorem," *Journal of Time Series Analysis* **15**, 269-278.

[4] Beveridge (1925), "Weather and Harvest Cycle", *Economic Journal* **31**, 429-452.

[5] Bollerslev,T. and H. O. Mikkelson (1996), "Modeling and Pricing Long Memory in Stock Market Volatility," *Journal of Econometics* **73**, 151-184.

[6] Box, G. E. P. and G. M. Jenkins (1976), Time Series Analysis: Forecasting and Control, Revised Ed. San Francisco: Holden Day.

[7] Box, G. E. P. and G. C.Tiao (1975), "Intervention Analysis with Applications to Economic and Environmental Problems," *Journal of American Statistical Association* **70,** 70-79.

[8] Campbell, J.Y. and N.G. Mankiw (1987), "Are Output Fluctuations Transitory?," *Quarterly Journal of Economics* **102**, 857-880.

[9] Cati, R.C., M. G. P. Garcia and P. Perron (1999), "Unit Roots in the Presence of Abrupt Governmental Interventions with an Application to Brazilian Data," *Journal of Applied Econometrics* **14,** 27-56.

[10] Chatfield, C. (1996), The Analysis of Time Series: An Introduction, Chapman&Hall New York.

[11] Chen, C and L. Liu (1993), "Joint Estimation of Model Parameters and Outliers Effects in Time Series," *Journal of the American Statistical Association* **74**, 427-431

[12] Cheung, Y.W. (1993), "Long Memory in Foreign Exchange Rates," *Journal of Business and Economic Statistics* **11**, 93-101.

[13] Chung, C. (1996),"Estimating a Generalized Long Memory Process," *Journal of Econometrics* **73**, 237-259.

[14] Comte, F. and E. Renault (1996), "Long Memory Continuous Time Models," *Journal of Econometrics* **73**, 101-149.

[15] Crato, N. and P. Rothman (1994), "Fractional Integration Analysis of Long-Run Behavior for Macroeconomic Time Series," *Economic Letters* **45**, 287-291.

[16] Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* **74**, 427-431

[17] Diebold, F.X. (1989), "Random Walks Versus Fractional Integration : Power Comparisons of Scaler and Joint Tests of the Variance-Time Function" in B.Raj (eds.), Advsnces in Econometrics, Kluwer Academic Publishers, Needham, Ma. 29-45.

[18] Diebold, F.X. and G.D. Rudebusch (1989), "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics* **24**, 189-209.

[19] Ding, Z., C. W. J. Granger and R. F. Engle (1993), "A Long Memory Property of Stock Returns and a New Model," *Journal of Empirical Finance* **1**, 83-106.

[20] Ding, Z. and C. W. J. Granger (1996), "Modeling Volatility Persistence of Speculative Returns: A new Approach", *Journal of Econometrics* **73**, 185-215.

[21] Franses, P.H. and N. Haldrup (1994), "The Effects of Additive Outliers on Tests for Unit Roots and Cointegration", *Journal of Business & Economic Statistics* **12,** 471-478.

[22] Fox, A.J. (1972), "Outliers in Time Series", *Journal of the Royal Statistical Association* **74,** 140-146.

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

[24] Gradszteyn, I.S. and I.M. Ryzhnik (1980), Tables of Integrals, Series and Products, 4th ed., Academic Press, New York.

[25] Granger, C.W.J. (1966) "The Typical Spectral Shape of an Economic Variable," *Econometrica* **34**, 150-161.

[26] Granger, C. W. J. (1980), "Long Memory Relationships and the Aggregation of Dynamic Models", *Journal of Econometrics* **14**, 227-238.

[27] Granger, C. W. J. (1981), "Some Properties of Time Series Data and Their Use in Econometric Model Specification", *Journal of Econometrics* **16**, 121-130.

[28] Granger, C.W.J. (1988), "Models That Generate Trends," *Journal of Time Series Analysis* **9**, 329-343.

[29] Granger, C.W.J. and Z. Ding (1996), "Varieties of Long Memory Models," *Journal of Econometrics* **73**, 61-77.

[30] Granger, C. W. J. and R. Joyeux (1980), "An Introduction to Long Memory Time Series Models and Fractional Differencing ", *Jurnal of Time Series Analysis* **1**, 15-39.

[31] Hawkins, D.M. (1980), Identification of Outliers, New York : Chapman and Hall.

[32] Hosking, J. R. M. (1981), "Fractional Differencing", *Biometrika* **68**, 165-176.

[33] Hosking, J. R. M. (1996), "Asymptotic Distributions of the Sample Mean, Autocovariances, and Autocorrelations of Long Memory Time Series", *Journal of Econometrics* **73**, 261-284.

[34] Hosoya, Y. (1996), "The Quasi-Likelihood Approach to Statistical Inference on Multiple Time Series with Long Range Dependence", *Journal of Econometrics* **73**, 217-236.

[35] Hurst, H. E. (1951), "Long Term Storage Capacity of Reserviors", *Transactions of the American Society of Civil Engineers* **116**, 770-799.

[36] Kennedy, P. (1998), A Guide to Econometrics, 4th ed., The MIT Press.

[37] Kokoszka, P.S. and M.S. Taqqu (1996), "Infinite Variance Stable Moving Averages with Long Memory," *Journal of Econometrics* **73**, 79-99.

[38] Li, W. K. and A. I. McLeod (1986), "Fractional Time Series Modeling", *Biometrika* **73**, 217-221.

[39] Lee, D. and P. Schmidt (1996), "On the Power of the KPSS Ttest of Stationarity against Fractionality Integrated Alternatives", *Journal of Econometrics* **73**, 285-302.

[40] Lo, A.W. (1991), "Long Term Memory in Stock Market Prices", *Econometrica* **59**, 1279-1313.

[41] Lobato, I. and P.M. Robinson (1996), "Averaged Periodogram Estimation of Long Term Memory", *Journal of Econometrics* **73**, 303-324.

[42] Mandelbrot, B.B. (1972), "Statistical Methodology for Non-Periodic Cycles : From the Covariance to R/S Analysis," *Annals of Economi and Social Management* **1**, 259-290.

[43] Mandelbrot, B.B. (1975), "A Fast Fractional Gaussian Noise Generator," *Water Resources Research* **7**, 543-553.

[44] McLeod, A.I. and K.W.Hipel (1978), "Preservation of the Rescaled Adjusted Range: A Reassessment of the Hurst Phenomenon", *Water Resources Research* **14**, 491-508.

[45] Peña, D. (1990), "Influential Observations in Time Series," *Journal of Business & Economic Statistics* **8** , 235-241.

[46] Perron, P. (1989), "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis", *Econometrica* **57,** 1361-1401.

[47] Perron, P. and S. Ng (1996), "Useful Modifications to Unit Root Tests with Dependent Errors and their Local Asymptotic Properties", *Review of Economic Studies* **63,** 435-463.

[48] Perron, P. and G. Rodríguez (2000), "Searching for Additive Outliers in Nonstationary Time Series", Working Paper **0005E,** Department of Economics, University of Ottawa.

[49] Robinson, P.M. (1990), Time Series with Strong Dependence, Advances in Econometrics, 6th world Congress, Cambridge University Press, Cambridge.

[50] Rodríguez, G. (2000), "An Empirical Note About Additive Outliers and Non-Stationarity in Latin American Inflation Series," A manuscript.

[51] Shea, G.S. (1991), "Uncertainty and Implied Variance Bounds in Long Memory Models of the Interest Rate Term Structure," *Empirical Economics* **16** 287-312.

[52] Sowell, F.B. (1990), "The Fractional Unit Root Distribution," *Econometrica* **58,** 495-505.

[53] Sowell, F.B. (1992), "Modeling Long Run Behavior with the Fractional ARIMA Model", *Journal of Monetary Economics* **29**, 277-302.

[54] Stock, J.H. (1990), "A Class of Tests for Integration and Cointegration", unpublished manuscript, Kennedy School of Government, Harvard University.

[55] Taqqu, M.S. (1975), "Weak Convergence to Fractional Brownian Motion and to the Rosenblatt Process", *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete* **31**, 287-302.

[56] Vogelsang, T.J. (1999), "Two Simple Procedures for Testing for a Unit Root when There are Additive Outliers", *Journal of Time Series Analysis* **20,** 237-252.

Table 1a. Bias, MSE of parameter d and size of t-statistic, $T = 50$
$\delta_1 = 0; \delta_2 = 0; \delta_3 = 0; \delta_4 = 0$

Parameter d	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.18	0.20	0.61	0.96
-0.72	0.09	0.17	0.46	0.95
-0.48	0.05	0.16	0.30	0.92
-0.24	0.03	0.16	0.16	0.85
0.00	0.02	0.16	0.10	0.75
0.24	0.02	0.16	0.20	0.56
0.48	0.03	0.16	0.38	0.35
0.72	0.05	0.15	0.60	0.17
0.96	0.02	0.13	0.82	0.11

Table1b Bias, MSE of parameter d and size of t-statistic, $T = 100$,
$\delta_1 = 0; \delta_2 = 0; \delta_3 = 0; \delta_4 = 0$

Parameter d	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.21	0.14	0.74	1.00
-0.72	0.09	0.10	0.62	1.00
-0.48	0.02	0.09	0.40	0.99
-0.24	0.01	0.09	0.20	0.97
0.00	0.01	0.09	0.08	0.90
0.24	0.01	0.09	0.19	0.73
0.48	0.02	0.09	0.46	0.46
0.72	0.03	0.09	0.76	0.20
0.96	0.02	0.07	0.92	0.09

Table 1c. Bias, MSE of parameter d and size of t-statistic, $T = 200$
$\delta_1 = 0; \delta_2 = 0; \delta_3 = 0; \delta_4 = 0$

Parameter d	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.20	0.13	0.86	1.00
-0.72	0.07	0.06	0.80	1.00
-0.48	0.02	0.05	0.56	1.00
-0.24	0.00	0.05	0.22	1.00
0.00	0.00	0.05	0.06	0.98
0.24	0.00	0.05	0.27	0.90
0.48	0.01	0.05	0.62	0.64
0.72	0.03	0.06	0.89	0.24
0.96	0.01	0.05	0.97	0.07

Table 1d. Bias, MSE of parameter d and size of t-statistic, $T = 500$
$\delta_1 = 0; \delta_2 = 0; \delta_3 = 0; \delta_4 = 0$

Parameter d	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.24	0.12	1.00	1.00
-0.72	0.09	0.04	1.00	1.00
-0.48	0.03	0.03	1.00	1.00
-0.24	0.01	0.03	1.00	1.00
0.00	0.01	0.03	0.05	1.00
0.24	0.01	0.03	1.00	1.00
0.48	0.02	0.03	1.00	1.00
0.72	0.04	0.03	1.00	1.00
0.96	0.01	0.03	1.00	0.09

Table2a. Bias, MSE of parameter d and size of t-statistic, $T = 50$
$\delta_1 = 10; \delta_2 = 5; \delta_3 = 2; \delta_4 = 2$

Parameter d	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.86	0.76	0.03	0.95
-0.72	0.61	0.41	0.04	0.95
-0.48	0.37	0.18	0.04	0.93
-0.24	0.13	0.08	0.06	0.92
0.00	-0.09	0.11	0.07	0.86
0.24	-0.23	0.18	0.08	0.77
0.48	-0.27	0.19	0.13	0.58
0.72	-0.24	0.19	0.35	0.34
0.96	-0.19	0.16	0.63	0.15

Table2b. Bias, MSE of parameter d and size of t-statistic, $T = 100$
$\delta_1 = 10; \delta_2 = 5; \delta_3 = 2; \delta_4 = 2$

Parameter d	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.82	0.68	0.05	0.99
-0.72	0.57	0.34	0.06	0.99
-0.48	0.32	0.13	0.07	0.98
-0.24	0.07	0.05	0.10	0.98
0.00	-0.11	0.08	0.09	0.96
0.24	-0.18	0.11	0.08	0.88
0.48	-0.17	0.11	0.26	0.68
0.72	-0.12	0.11	0.63	0.33
0.96	-0.08	0.08	0.87	0.12

Table 2c. Bias, MSE of parameter d and size of t-statistic, $T = 200$
$\delta_1 = 10; \delta_2 = 5; \delta_3 = 2; \delta_4 = 2$

Parameter d	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.93	0.88	0.01	1.00
-0.72	0.67	0.46	0.02	1.00
-0.48	0.40	0.18	0.03	0.99
-0.24	0.14	0.06	0.07	0.99
0.00	-0.01	0.05	0.07	0.98
0.24	-0.05	0.06	0.19	0.92
0.48	-0.03	0.06	0.56	0.69
0.72	0.00	0.06	0.87	0.28
0.96	0.00	0.05	0.97	0.07

Table2d. Bias, MSE of parameter d and size of t-statistic, $T = 500$
$\delta_1 = 10; \delta_2 = 5; \delta_3 = 2; \delta_4 = 2$

Parameter d	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.91	0.84	1.00	1.00
-0.72	0.65	0.43	1.00	1.00
-0.48	0.36	0.14	1.00	1.00
-0.24	0.09	0.03	0.96	1.00
0.00	-0.01	0.03	0.05	1.00
0.24	-0.01	0.03	0.98	1.00
0.48	0.01	0.03	1.00	1.00
0.72	0.03	0.03	1.00	1.00
0.96	0.01	0.03	1.00	0.63

Table 3a. Bias, MSE of parameter d and size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 50$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.3	0.57	0.45	0.26	0.94
	0.4	0.60	0.48	0.23	0.93
	0.5	0.60	0.49	0.25	0.95
	0.6	0.58	0.46	0.24	0.95
	0.7	0.57	0.44	0.26	0.95
	0.8	0.52	0.40	0.30	0.95
	0.9	0.42	031	0.40	0.98
-0.72	0.3	0.35	0.25	0.25	0.94
	0.4	0.38	0.26	0.23	0.93
	0.5	0.38	0.27	0.23	0.93
	0.6	0.37	0.26	0.25	0.94
	0.7	0.35	0.25	0.26	0.95
	0.8	0.32	0.22	0.25	0.95
	0.9	0.25	0.19	0.35	0.97

Table 3a.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are Additive Outliersgenerated according to a Bernoulli Distribution, $T = 50$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.48	0.3	0.18	0.16	0.21	0.94
	0.4	0.20	0.15	0.19	0.93
	0.5	0.20	0.16	0.20	0.92
	0.6	0.19	0.18	0.19	0.92
	0.7	0.16	0.16	0.21	0.93
	0.8	0.16	0.14	0.22	0.94
	0.9	0.11	0.14	0.26	0.95
-0.24	0.3	0.06	0.12	0.16	0.88
	0.4	0.06	0.13	0.12	0.90
	0.5	0.07	0.13	0.15	0.89
	0.6	0.08	0.14	0.13	0.87
	0.7	0.04	0.13	0.14	0.89
	0.8	0.06	0.12	0.14	0.89
	0.9	0.04	0.12	0.15	0.91

Table 3a.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are Additive Outliersgenerated according to a Bernoulli Distribution, $T = 50$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.00	0.3	-0.01	0.13	0.09	0.81
	0.4	-0.01	0.12	0.09	0.79
	0.5	-0.01	0.13	0.09	0.80
	0.6	-0.01	0.13	0.10	0.82
	0.7	-0.02	0.13	0.11	0.80
	0.8	-0.01	0.12	0.08	0.80
	0.9	-0	0.12	0.10	0.82
0.24	0.3	-0.04	0.14	0.14	0.66
	0.4	-0.03	0.12	0.14	0.62
	0.5	-0.04	0.12	0.16	0.66
	0.6	-0.02	0.13	0.15	0.64
	0.7	-0.04	0.13	0.17	0.65
	0.8	0	0.13	0.17	0.64
	0.9	-0.02	0.12	0.17	0.62

Table 3a.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are Additive Outliersgenerated according to a Bernoulli Distribution, $T = 50$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.48	0.3	-0.04	0.13	0.36	0.43
	0.4	-0.02	0.12	0.37	0.42
	0.5	-0.04	0.12	0.34	0.42
	0.6	-0.02	0.13	0.36	0.42
	0.7	-0.03	0.13	0.37	0.41
	0.8	0.00	0.13	0.37	0.41
	0.9	-0.01	0.12	0.39	0.39
0.72	0.3	-0.02	0.12	0.62	0.20
	0.4	0.00	0.12	0.61	0.22
	0.5	-0.02	0.13	0.62	0.21
	0.6	0.00	0.13	0.62	0.21
	0.7	-0.01	0.13	0.62	0.19
	0.8	0.01	0.13	0.63	0.20
	0.9	0.01	0.12	0.66	0.20

Table 3a.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are Additive Outliersgenerated according to a Bernoulli Distribution, $T = 50$

Parameter d	p	Bias	MSE	Size of t-statisitic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.96	0.3	-0.03	0.11	0.85	0.09
	0.4	-0.01	0.10	0.84	0.10
	0.5	-0.03	0.11	0.83	0.09
	0.6	-0.00	0.11	0.84	0.10
	0.7	-0.02	0.11	0.84	0.09
	0.8	0.01	0.12	0.84	0.11
	0.9	-0.02	0.11	0.86	0.10

Table 3b. Bias, MSE of parameter d and Size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 100$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.3	0.70	0.56	0.20	0.99
	0.4	0.72	0.58	0.21	0.99
	0.5	0.72	0.60	0.19	0.99
	0.6	0.70	0.56	0.18	0.98
	0.7	0.71	0.57	0.20	0.99
	0.8	0.66	0.51	0.26	0.99
	0.9	0.56	0.39	0.38	1.00
-0.72	0.3	0.44	0.27	0.25	0.99
	0.4	0.47	0.82	0.22	0.98
	0.5	0.47	0.29	0.21	0.99
	0.6	0.45	0.27	0.20	0.99
	0.7	0.46	0.27	0.24	0.99
	0.8	0.41	0.24	0.24	0.99
	0.9	0.33	0.18	0.37	0.99

Table 3b.(Cont'd.) Bias, MSE of parameter d and Size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 100$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.48	0.3	0.23	0.13	0.19	0.99
	0.4	0.24	0.12	0.20	0.99
	0.5	0.24	0.13	0.20	0.99
	0.6	0.22	0.12	0.20	0.99
	0.7	0.23	0.12	0.20	0.99
	0.8	0.20	0.12	0.25	0.99
	0.9	0.15	0.09	0.29	0.99
-0.24	0.3	0.08	0.08	0.15	0.98
	0.4	0.08	0.08	0.12	0.98
	0.5	0.08	0.07	0.12	0.99
	0.6	0.06	0.07	0.13	0.98
	0.7	0.08	0.07	0.13	0.98
	0.8	0.07	0.08	0.13	0.98
	0.9	0.04	0.07	0.16	0.99

Table 3b.(Cont'd.) Bias, MSE of parameter d and Size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 100$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0	0.3	0.01	0.07	0.08	0.96
	0.4	0.00	0.07	0.08	0.95
	0.5	-0.01	0.07	0.07	0.95
	0.6	-0.02	0.07	0.08	0.96
	0.7	0.00	0.07	0.09	0.95
	0.8	0.01	0.07	0.09	0.94
	0.9	0.00	0.07	0.08	0.94
0.24	0.3	-0.01	0.06	0.17	0.83
	0.4	-0.03	0.07	0.20	0.85
	0.5	-0.03	0.07	0.18	0.85
	0.6	-0.05	0.07	0.19	0.84
	0.7	-0.01	0.07	0.19	0.85
	0.8	-0.01	0.07	0.21	0.83
	0.9	-0.01	0.07	0.23	0.84

Table 3b.(Cont'd.) Bias, MSE of parameter d and Size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 100$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.48	0.3	-0.01	0.07	0.54	0.55
	0.4	-0.01	0.07	0.52	0.58
	0.5	-0.02	0.06	0.51	0.57
	0.6	-0.04	0.07	0.52	0.57
	0.7	-0.01	0.07	0.52	0.55
	0.8	-0.01	0.07	0.53	0.55
	0.9	0.00	0.07	0.52	0.53
0.72	0.3	0.01	0.06	0.80	0.24
	0.4	0.01	0.07	0.79	0.22
	0.5	0.01	0.06	0.81	0.25
	0.6	-0.01	0.07	0.81	0.24
	0.7	0.01	0.07	0.82	0.24
	0.8	0.01	0.07	0.81	0.22
	0.9	0.01	0.07	0.83	0.21

Table 3b.(Cont'd.) Bias, MSE of parameter d and Size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 100$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.96	0.3	-0.06	0.06	0.96	0.08
	0.4	0.00	0.06	0.96	0.08
	0.5	0.00	0.05	0.96	0.08
	0.6	-0.01	0.06	0.95	0.08
	0.7	-0.01	0.07	0.96	0.08
	0.8	0.01	0.05	0.96	0.07
	0.9	-0.01	0.06	0.96	0.08

Table 3c. Bias, MSE of parameter d and size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 200$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	0.3	0.79	0.66	0.17	1.00
	0.4	0.80	0.68	0.16	1.00
	0.5	0.81	0.69	0.15	1.00
	0.6	0.81	0.69	0.16	1.00
	0.7	0.80	0.68	0.16	1.00
	0.8	0.75	0.61	0.21	1.00
	0.9	0.69	0.51	0.33	1.00
-0.72	0.3	0.52	0.31	0.22	1.00
	0.4	0.53	0.32	0.21	1.00
	0.5	0.54	0.33	0.17	1.00
	0.6	0.53	0.31	0.19	1.00
	0.7	0.52	0.27	0.20	1.00
	0.8	0.48	0.21	0.28	1.00
	0.9	0.41	0.21	0.37	1.00

Table 3c.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 200$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.48	0.3	0.21	0.11	0.23	1.00
	0.4	0.28	0.12	0.20	1.00
	0.5	0.29	0.12	0.20	1.00
	0.6	0.28	0.12	0.20	1.00
	0.7	0.27	0.11	0.22	1.00
	0.8	0.24	0.09	0.27	1.00
	0.9	0.19	0.07	0.37	1.00
-0.24	0.3	0.09	0.05	0.16	1.00
	0.4	0.09	0.05	0.16	1.00
	0.5	0.10	0.05	0.14	1.00
	0.6	0.08	0.04	0.13	1.00
	0.7	0.09	0.04	0.16	1.00
	0.8	0.07	0.04	0.17	1.00
	0.9	0.05	0.04	0.20	1.00

Table 3c.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 200$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0	0.3	0.01	0.04	0.07	1.00
	0.4	0	0.04	0.08	1.00
	0.5	0	0.04	0.08	1.00
	0.6	-0.02	0.04	0.06	1.00
	0.7	0.00	0.04	0.05	1.00
	0.8	0.00	0.04	0.07	0.99
	0.9	0.00	0.04	0.07	1.00
0.24	0.3	-0.01	0.04	0.23	0.97
	0.4	0.02	0.04	0.24	0.97
	0.5	-0.02	0.04	0.23	0.96
	0.6	-0.04	0.04	0.25	0.96
	0.7	-0.02	0.04	0.24	0.96
	0.8	-0.01	0.04	0.26	0.96
	0.9	-0.01	0.04	0.26	0.96

Table 3c.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 200$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.48	0.3	0.00	0.04	0.72	0.77
	0.4	-0.01	0.04	0.71	0.77
	0.5	0.00	0.04	0.71	0.79
	0.6	-0.03	0.04	0.71	0.80
	0.7	-0.01	0.04	0.71	0.79
	0.8	0.02	0.04	0.71	0.78
	0.9	0.00	0.04	0.72	0.77
0.72	0.3	0.03	0.04	0.95	0.33
	0.4	0.01	0.02	0.93	0.31
	0.5	0.02	0.04	0.94	0.34
	0.6	0.00	0.04	0.95	0.31
	0.7	0.02	0.04	0.95	0.31
	0.8	0.02	0.04	0.95	0.30
	0.9	0.02	0.04	0.95	0.32

Table 3c.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are Additive Outliers generated according to a Bernoulli Distribution, $T = 200$

Parameter d	p	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.96	0.3	0.01	0.03	0.99	0.08
	0.4	0.00	0.03	0.99	0.07
	0.5	0.01	0.04	0.99	0.07
	0.6	-0.01	0.03	0.99	0.06
	0.7	0.00	0.03	0.99	0.06
	0.8	0.01	0.03	0.99	0.06
	0.9	0.01	0.03	0.99	0.06

Table 4a. Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 50$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	-0.8	0	1.12	1.41	0.11	0.65
	-0.8	4	1.14	1.46	0.11	0.65
	-0.4	0	1.15	1.27	0.11	0.69
	-0.4	4	1.04	1.25	0.11	0.69
	0.0	0	1.01	1.17	0.12	0.72
	0.0	4	1.01	1.18	0.12	0.72
	0.4	0	1.03	1.22	0.11	0.71
	0.4	4	1.03	1.24	0.11	0.71
	0.8	0	1.15	1.48	0.10	0.66
	0.8	4	1.11	1.39	0.10	0.66
-0.72	-0.8	0	0.88	0.94	0.10	0.63
	-0.8	4	0.91	0.98	0.10	0.63
	-0.4	0	0.82	0.84	0.12	0.67
	-0.4	4	0.82	0.83	0.12	0.67
	0.0	0	0.79	0.77	0.11	0.70
	0.0	4	0.79	0.79	0.11	0.70
	0.4	0	0.81	0.82	0.10	0.69
	0.4	4	0.81	0.82	0.10	0.69
	0.8	0	0.92	1.00	0.09	0.65
	0.8	4	0.87	0.92	0.09	0.65

Table 4a.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 50$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.48	-0.8	0	0.64	0.57	0.11	0.64
	-0.8	4	0.67	0.59	0.11	0.64
	-0.4	0	0.60	0.52	0.11	0.65
	-0.4	4	0.60	0.51	0.11	0.65
	0.0	0	0.57	0.48	0.11	0.66
	0.0	4	0.58	0.50	0.11	0.66
	0.4	0	0.59	0.51	0.10	0.68
	0.4	4	0.59	0.51	0.09	0.66
	0.8	0	0.68	0.60	0.09	0.66
	0.8	4	0.63	0.54	0.11	0.65
-0.24	-0.8	0	0.40	0.31	0.11	0.64
	-0.8	4	0.43	0.32	0.12	0.64
	-0.4	0	0.39	0.30	0.12	0.64
	-0.4	4	0.38	0.30	0.12	0.65
	0.0	0	0.38	0.29	0.12	0.65
	0.0	4	0.39	0.31	0.12	0.64
	0.4	0	0.39	0.30	0.12	0.64
	0.4	4	0.39	0.31	0.12	0.64
	0.8	0	0.43	0.32	0.12	0.64
	0.8	4	0.39	0.29	0.12	0.64

Table 4a.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 50$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0	-0.8	0	0.19	0.18	0.15	0.63
	-0.8	4	0.21	0.18	0.15	0.63
	-0.4	0	0.20	0.17	0.15	0.63
	-0.4	4	0.19	0.19	0.15	0.63
	0.0	0	0.20	0.19	0.15	0.63
	0.0	4	0.21	0.18	0.15	0.63
	0.4	0	0.20	0.18	0.15	0.63
	0.4	4	0.20	0.18	0.15	0.63
	0.8	0	0.21	0.19	0.15	0.63
	0.8	4	0.18	0.17	0.15	0.63
0.24	-0.8	0	0.06	0.13	0.17	0.53
	-0.8	4	0.08	0.12	0.17	0.53
	-0.4	0	0.04	0.13	0.19	0.55
	-0.4	4	0.03	0.15	0.19	0.55
	0.0	0	0.04	0.15	0.21	0.57
	0.0	4	0.05	0.14	0.21	0.57
	0.4	0	0.05	0.14	0.20	0.58
	0.4	4	0.05	0.13	0.20	0.58
	0.8	0	0.06	0.14	0.18	0.54
	0.8	4	0.06	0.12	0.18	0.54

Table 4a.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 50$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.48	-0.8	0	-0.12	0.09	0.07	0.35
	-0.8	4	-0.12	0.10	0.07	0.35
	-0.4	0	-0.11	0.13	0.19	0.46
	-0.4	4	-0.13	0.14	0.19	0.46
	0.0	0	-0.08	0.16	0.30	0.46
	0.0	4	-0.08	0.14	0.30	0.46
	0.4	0	-0.12	0.14	0.21	0.44
	0.4	4	-0.12	0.14	0.21	0.44
	0.8	0	-0.12	0.10	0.06	0.32
	0.8	4	-0.13	0.11	0.06	0.32
0.72	-0.8	0	-0.42	0.22	0.00	0.28
	-0.8	4	-0.42	0.22	0.00	0.28
	-0.4	0	-0.35	0.22	0.11	0.34
	-0.4	4	-0.36	0.23	0.11	0.34
	0.0	0	-0.21	0.18	0.38	0.32
	0.0	4	-0.21	0.17	0.38	0.32
	0.4	0	-0.35	0.22	0.10	0.31
	0.4	4	-0.39	0.22	0.10	0.31
	0.8	0	-0.42	0.22	0.00	0.6
	0.8	4	-0.44	0.23	0.00	0.26

Table 4a.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 50$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.96	-0.8	0	-0.72	0.53	0.00	0.32
	-0.8	4	-0.73	0.54	0.00	0.32
	-0.4	0	-0.65	0.47	0.04	0.34
	-0.4	4	-0.66	0.49	0.04	0.34
	0.0	0	-0.39	0.28	0.38	0.25
	0.0	4	-0.39	0.27	0.38	0.25
	0.4	0	-0.64	0.46	0.03	0.33
	0.4	4	-0.65	0.47	0.03	0.33
	0.8	0	-0.72	0.53	0.00	0.32
	0.8	4	-0.73	0.55	0.00	0.32

Table 4b. Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 100$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	-0.8	0	0.94	0.96	0.07	0.94
	-0.8	4	0.94	0.96	0.07	0.94
	-0.4	0	0.89	0.88	0.09	0.95
	-0.4	4	0.89	0.88	0.09	0.95
	0.0	0	0.88	0.85	0.10	0.96
	0.0	4	0.88	0.85	0.10	0.96
	0.4	0	0.89	0.87	0.09	0.96
	0.4	4	0.89	0.87	0.09	0.96
	0.8	0	0.93	0.94	0.07	0.95
	0.8	4	0.93	0.94	0.07	0.95
-0.72	-0.8	0	0.69	0.55	0.07	0.95
	-0.8	4	0.69	0.55	0.07	0.95
	-0.4	0	0.65	0.49	0.09	0.95
	-0.4	4	0.65	0.49	0.09	0.95
	0.0	0	0.64	0.48	0.10	0.95
	0.0	4	0.64	0.48	0.10	0.95
	0.4	0	0.65	0.49	0.10	0.95
	0.4	4	0.65	0.49	0.10	0.95
	0.8	0	0.68	0.54	0.08	0.95
	0.8	4	0.68	0.54	0.08	0.95

Table 4b.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 100$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.48	-0.8	0	0.44	0.26	0.07	0.95
	-0.8	4	0.44	0.26	0.07	0.95
	-0.4	0	0.42	0.24	0.08	0.96
	-0.4	4	0.42	0.24	0.08	0.96
	0.0	0	0.41	0.24	0.10	0.96
	0.0	4	0.41	0.24	0.10	0.96
	0.4	0	0.42	0.25	0.09	0.95
	0.4	4	0.42	0.25	0.09	0.95
	0.8	0	0.44	0.26	0.08	0.96
	0.8	4	0.44	0.26	0.08	0.96
-0.24	-0.8	0	0.22	0.11	0.06	0.94
	-0.8	4	0.22	0.11	0.06	0.94
	-0.4	0	0.21	0.11	0.07	0.95
	-0.4	4	0.21	0.11	0.07	0.95
	0.0	0	0.22	0.11	0.07	0.96
	0.0	4	0.22	0.11	0.07	0.96
	0.4	0	0.22	0.11	0.07	0.94
	0.4	4	0.22	0.11	0.07	0.94
	0.8	0	0.22	0.12	0.08	0.94
	0.8	4	0.22	0.12	0.08	0.94

Table 4b.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 100$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0	-0.8	0	0.06	0.06	0.07	0.93
	-0.8	4	0.06	0.06	0.07	0.93
	-0.4	0	0.06	0.06	0.07	0.93
	-0.4	4	0.06	0.06	0.07	0.93
	0.0	0	0.06	0.06	0.07	0.93
	0,0	4	0.06	0.06	0.07	0.93
	0.4	0	0.06	0.06	0.07	0,93
	0.4	4	0.06	0.06	0.07	0.93
	0.8	0	0.06	0.06	0.07	0.93
	0.8	4	0.06	0.06	0.07	0.93
0.24	-0.8	0	-0.09	0.09	0.13	0.87
	-0.8	4	-0.09	0.09	0.13	0.87
	-0.4	0	-0.06	0.08	0.16	0.83
	-0.4	4	-0.06	0.08	0.16	0.83
	0.0	0	-0.04	0.07	0.16	0.85
	0.0	4	-0.04	0.07	0.16	0.85
	0.4	0	-0.06	0.07	0.16	0.88
	0.4	4	-0.06	0.07	0.16	0.88
	0.8	0	-0.10	0.09	0.11	0.88
	0.8	4	-0.10	0.09	0.11	0.88

Table 4b.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers,$T = 100$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.48	-0.8	0	-0.35	0.16	0.04	0.95
	-0.8	4	-0.35	0.16	0.04	0.95
	-0.4	0	-0.25	0.12	0.17	0.83
	-0.4	4	-0.25	0.12	0.17	0.83
	0.0	0	-0.10	0.08	0.36	0.68
	0.0	4	-0.10	0.08	0.36	0.68
	0.4	0	-0.26	0.14	0.17	0.83
	0.4	4	-0.26	0.14	0.17	0.83
	0.8	0	-0.36	0.17	0.04	0.95
	0.8	4	-0.36	0.17	0.04	0.95
0.72	-0.8	0	-0.54	0.30	0.00	1.00
	-0.8	4	-0.54	0.30	0.00	1.00
	-0.4	0	-0.51	0.29	0.09	0.93
	-0.4	4	-0.51	0.29	0.09	0.93
	0.0	0	-0.23	0.12	0.52	0.51
	0.0	4	-0.23	0.12	0.52	0.51
	0.4	0	-0.51	0.30	0.07	0.93
	0.4	4	-0.51	0.30	0.07	0.93
	0.8	0	-0.54	0.30	0.01	1.00
	0.8	4	-0.54	0.30	0.01	1.00

Table 4b.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 100$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.96	-0.8	0	-0.72	0.53	0.00	1.00
	-0.8	4	-0.72	0.53	0.00	1.00
	-0.4	0	-0.72	0.53	0.03	0.99
	-0.4	4	-0.72	0.53	0.03	0.99
	0.0	0	-0.46	0.28	0.51	0.52
	0.0	4	-0.46	0.28	0.51	0.52
	0.4	0	-0.72	0.54	0.03	0.99
	0.4	4	-0.72	0.54	0.03	0.99
	0.8	0	-0.72	0.53	0.00	1.00
	0.8	4	-0.72	0.53	0.00	1.00

Table 4c. Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 200$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.96	-0.8	0	0.87	0.82	0.12	1.00
	-0.8	4	0.87	0.82	0.12	1.00
	-0.4	0	0.83	0.75	0.15	1.00
	-0.4	4	0.83	0.75	0.15	1.00
	0.0	0	0.81	0.71	0.18	1.00
	0.0	4	0.81	0.71	0.18	1.00
	0.4	0	0.82	0.73	0.17	1.00
	0.4	4	0.82	0.73	0.17	1.00
	0.8	0	0.86	0.79	0.14	1.00
	0.8	4	0.86	0.79	0.14	1.00
-0.72	-0.8	0	0.61	0.42	0.12	1.00
	-0.8	4	0.61	0.42	0.12	1.00
	-0.4	0	0.57	0.37	0.18	1.00
	-0.4	4	0.57	0.37	0.18	1.00
	0.0	0	0.54	0.35	0.23	1.00
	0.0	4	0.54	0.35	0.23	1.00
	0.4	0	0.56	0.36	0.18	1.00
	0.4	4	0.56	0.36	0.18	1.00
	0.8	0	0.60	0.40	0.16	1.00
	0.8	4	0.60	0.40	0.16	1.00

Table 4c. (Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 200$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
-0.48	-0.8	0	0.36	0.17	0.12	1.00
	-0.8	4	0.36	0.17	0.12	1.00
	-0.4	0	0.33	0.15	0.18	1.00
	-0.4	4	0.33	0.15	0.18	1.00
	0.0	0	0.31	0.15	0.20	1.00
	0.0	4	0.31	0.15	0.20	1.00
	0.4	0	0.32	0.15	0.17	1.00
	0.4	4	0.32	0.15	0.17	1.00
	0.8	0	0.36	0.17	0.12	1.00
	0.8	4	0.36	0.17	0.12	1.00
-0.24	-0.8	0	0.15	0.06	0.10	1.00
	-0.8	4	0.15	0.06	0.10	1.00
	-0.4	0	0.13	0.06	0.12	1.00
	-0.4	4	0.13	0.06	0.12	1.00
	0.0	0	0.13	0.06	0.12	1.00
	0.0	4	0.13	0.06	0.12	1.00
	0.4	0	0.13	0.06	0.12	1.00
	0.4	4	0.13	0.06	0.12	1.00
	0.8	0	0.15	0.06	0.09	1.00
	0.8	4	0.15	0.06	0.09	1.00

Table 4c. (Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 200$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0	-0.8	0	0.02	0.04	0.06	0.99
	-0.8	4	0.02	0.04	0.06	0.99
	-0.4	0	0.02	0.04	0.06	0.99
	-0.4	4	0.02	0.04	0.06	0.99
	0.0	0	0.02	0.04	0.06	0.99
	0.0	4	0.02	0.04	0.06	0.99
	0.4	0	0.02	0.04	0.06	0.99
	0.4	4	0.02	0.04	0.06	0.99
	0.8	0	0.02	0.04	0.06	0.99
	0.8	4	0.02	0.04	0.06	0.99
0.24	-0.8	0	0.02	0.04	0.31	0.95
	-0.8	4	0.02	0.04	0.31	0.95
	-0.4	0	-0.01	0.04	0.26	0.96
	-0.4	4	-0.01	0.04	0.26	0.96
	0.0	0	-0.03	0.04	0.22	0.96
	0.0	4	-0.03	0.04	0.22	0.96
	0.4	0	-0.01	0.04	0.25	0.96
	0.4	4	-0.01	0.04	0.25	0.96
	0.8	0	0.02	0.04	0.31	0.96
	0.8	4	0.02	0.04	0.31	0.96

Table 4c. (Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 200$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.48	-0.8	0	-0.01	0.01	0.79	0.86
	-0.8	4.	-0.01	0.01	0.79	0.86
	-0.4	0	-0.03	0.03	0.66	0.82
	-0.4	4	-0.03	0.03	0.66	0.82
	0.0	0	-0.06	0.04	0.60	0.83
	0.0	4	-0.06	0.04	0.60	0.83
	0.4	0	-0.04	0.03	0.65	0.83
	0.4	4	-0.04	0.03	0.65	0.83
	0.8	0	-0.01	0.01	0.79	0.85
	0.8	4	-0.01	0.01	0.79	0.85
0.72	-0.8	0	-0.16	0.03	0.97	0.69
	-0.8	4	-0.16	0.03	0.97	0.69
	-0.4	0	-0.18	0.04	0.92	0.72
	-0.4	4	-0.18	0.04	0.92	0.72
	0.0	0	-0.12	0.05	0.87	0.56
	0.0	4	-0.12	0.05	0.87	0.56
	0.4	0	-0.18	0.04	0.91	0.71
	0.4	4	-0.18	0.04	0.91	0.71
	0.8	0	-0.16	0.03	0.97	0.69
	0.8	4	-0.16	0.03	0.97	0.69

Table 4c.(Cont'd.) Bias, MSE of parameter d and size of t-statistic when there are inliers, $T = 200$

Parameter d	μ	a	Bias	MSE	Size of t-statistic $H_0 : d = 0$	Size of t-statistic $H_0 : d = 1$
0.96	-0.8	0	-0.38	0.14	1.00	0.65
	-0.8	4	-.038	0.14	1.00	0.65
	-0.4	0	-0.38	0.15	0.97	0.58
	-0.4	4	-0.38	0.15	0.97	0.58
	0.0	0	-0.27	0.11	0.94	0.39
	0.0	4	-0.27	0.11	0.94	0.39
	0.4	0	-0.38	0.14	0.97	0.57
	0.4	4	-0.38	0.14	0.97	0.57
	0.8	0	-0.38	0.14	1.00	0.64
	0.8	4	-0.38	0.14	1.00	0.64