Estimating autocorrelations in the presence of deterministic trends

Shin-Huei Wang and Christian M. Hafner
CORE
Voie du Roman Pays 34
B-1348 Louvain-la-Neuve, Belgium.
Tel (32 10) 47 43 04
Fax (32 10) 47 43 01
E-mail: corestat-library@uclouvain.be
Estimating autocorrelations in the presence of deterministic trends

Shin-Huei WANG\textsuperscript{1} and Christian M. HAFNER\textsuperscript{2}

December 2008

Abstract

This paper considers the impact of ordinary least squares (OLS) detrending and the first difference (FD) detrending on autocorrelation estimation in the presence of long memory and deterministic trends. We show that the FD detrending results in inconsistent autocorrelation estimates when the error term is stationary. Thus, the FD detrending should not be employed for autocorrelation estimation of the detrended series when constructing e.g. portmanteau-type tests. In an empirical application of volume in Dow Jones stocks, we show that for some stocks, OLS and FD detrending result in substantial differences in ACF estimates.

Keywords: autocorrelations, OLS, first difference detrending, long memory.

JEL Classification: C22

\textsuperscript{1} Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium. E-mail: shin-huei.wang@uclouvain.be
\textsuperscript{2} Université catholique de Louvain, Institut de Statistique and CORE, B-1348 Louvain-la-Neuve, Belgium. E-mail: Christian.hafner@uclouvain.be

Financial support from the contract Projet d'Actions de Recherche Concertées nr. 07/12/002 of the Communauté française de Belgique, granted by the Académie universitaire Louvain, is gratefully acknowledged.

This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.
1 Introduction

This paper considers the problem of autocorrelation estimation in the linear trend model,

\[ y_t = \alpha + \beta t + v_t, \quad t = 1, \ldots, T \]  

(1)

where \( v_t \) follows a stationary ARFIMA(\( p, d, q \)) process with \( d \in (-0.5, 0.5) \). This model has been considered widely in the econometrics literature since many economic time series exhibit clear trends and long memory. For example, Yajima (1988, 1991) studies the efficiency of the OLS estimator under stationarity of \( v_t \). Lee and Phillips (1994) show that in the nonstationary case, efficiency gains can be obtained by generalized least squares. Canjels and Watson (1997) compare OLS with a simple first difference (FD) estimator and generalized least squares for the I(0) and I(1) cases. Tsay (2000) recommends the FD estimator under stationary or non-stationary long memory as its convergence rate is \( T^{-1} \), faster than the usual \( T^{-1/2} \). In the presence of additional stochastic regressors and non-stationarity, the FD estimator has a faster convergence rate than OLS for the coefficients of the stochastic regressors. This may explain the popularity of the FD estimator in econometric practice.

The focus of previous research, however, has mainly been on finding efficient estimators of \( \beta \). In this paper, we turn our attention to the autocorrelations of \( v_t \), defined as: \( \hat{\rho}_v(j) = \frac{\sum_{t=j+1}^{T} (v_t - \overline{v})(v_{t-j} - \overline{v})}{\sum_{t=1}^{T} (v_t - \overline{v})^2} \), for \( j = 0, 1, 2, \ldots \). These autocorrelations are useful in applications because they can be used as visual diagnostics of the characteristics of the observed time series. These can be used e.g. for model building, e.g. to determine the order \((p, d, q)\) of the ARFIMA model. They can also be used in constructing feasible generalized least squares estimators of the trend. Moreover, autocorrelations are the building blocks of the portmanteau-type tests, such as the Box-Pierce and the Ljung-Box test.

Empirically, we have observations \( y_t \) and cannot observe \( v_t \) directly. If we do not employ any detrending procedure and just use the original data \( y_t \) to calculate the sample autocorrelations, \( \hat{\rho}_y(j) = \frac{\sum_{t=j+1}^{T} (y_t - \overline{y})(y_{t-j} - \overline{y})}{\sum_{t=1}^{T} (y_t - \overline{y})^2} \), then the information contained in \( \hat{\rho}_y(j) \) will not be the one we need. In fact, we can prove that \( \hat{\rho}_y(j) \xrightarrow{p} 1 \) for any \( j \) and \( j = o(T) \), because the deterministic trend dominates the asymptotic behavior of \( y_t \).

To eliminate the influence of the deterministic trend on the autocorrelation estimation, two detrending methods are popularly used in the literature. The first method is the ordinary least squares (OLS) detrending, i.e., estimate (1) by OLS and obtain OLS
residuals, the second one is the first difference (FD) detrending, where $\beta$ is estimated by the sample mean of $\Delta y_t = y_t - y_{t-1}$. Both detrending methods do not depend on the magnitude of the differencing parameter $d$.

The FD estimator was proposed by Grenander and Rosenblatt (1957) in estimating the deterministic trending coefficient. It can be used as an approximation to the generalized least squares (GLS) estimator when high positive correlation of the error term is suspected, see e.g. Maeshiro (1976), Chipman (1979), Krämer (1982) and references therein. Furthermore, Tsay (2000) investigated the asymptotic properties of the FD estimator in the fractional cointegration model introduced by Granger (1986) and later investigated by Cheung and Lai (1993). Tsay (2000) showed that the FD estimator of the slope can result in faster convergence rates when the error term is nonstationary.

From the residuals of both OLS and FD detrending, we can construct two autocorrelation functions, $\hat{\rho}_{\text{OLS}}(j)$ and $\hat{\rho}_{\text{FD}}(j)$, respectively. We would like to show that $\hat{\rho}_{\text{OLS}}(j)$ and $\hat{\rho}_{\text{FD}}(j)$ are consistent estimators of their population counterpart, $\rho_v(j)$, i.e., the population autocorrelation function of $v_t$ at lag $j$. Our theoretical analysis reveals that $\hat{\rho}_{\text{OLS}}(j) \xrightarrow{p} \rho_v(j)$ as long as $v_t$ is stationary. On the other hand, the consistency of $\hat{\rho}_{\text{FD}}(j)$ does not hold, so that its use in empirical applications is not justified. A small set of Monte Carlo simulation results shows that the root mean squared error (RMSE) of $\hat{\rho}_{\text{OLS}}(j)$ decreases monotonically with the sample size. Furthermore, no matter what lag number and differencing parameter is used in the experiment, the RMSE of $\hat{\rho}_{\text{FD}}(j)$ does not diminish monotonically. This corroborates our inconsistency result of FD detrending. Therefore, the FD detrending method should not be employed for autocorrelation estimation of the detrended series.

2 Main results

In this section we consider the impact of OLS detrending and FD detrending on the autocorrelation estimation of $v_t$ in equation (1). The asymptotic properties of $\hat{\rho}_{\text{OLS}}(j)$ and $\hat{\rho}_{\text{FD}}(j)$ will be investigated when $v_t$ is a stationary $I(d)$ or ARFIMA process. Let us first introduce some notation and define the relevant statistics. OLS detrending uses the OLS estimators $\hat{\alpha}_{\text{OLS}}$ and $\hat{\beta}_{\text{OLS}}$ to obtain OLS residuals $e_{t,\text{OLS}} = y_t - \hat{\alpha}_{\text{OLS}} - \hat{\beta}_{\text{OLS}}t$. FD detrending uses the estimator $\hat{\beta}_{\text{FD}} = \frac{1}{T-1} \sum_{t=2}^{T} \Delta y_t$ to obtain residuals $e_{t,\text{FD}} = y_t - \hat{\beta}_{\text{FD}}t$. 

From these residuals, we construct two autocorrelation functions,

$$\hat{\rho}_{OLS}(j) = \frac{\sum_{t=j+1}^{T} e_{t,OLS} e_{t-j,OLS}}{\sum_{t=1}^{T} e_{t,OLS}^2}, \quad \hat{\rho}_{FD}(j) = \frac{\sum_{t=j+1}^{T} (e_{t,FD} - \bar{e}_{FD}) (e_{t-j,FD} - \bar{e}_{FD})}{\sum_{t=1}^{T} (e_{t,FD} - \bar{e}_{FD})^2},$$

where $\bar{e}_{FD}$ is the sample mean of $e_{t,FD}$, noting that the sample mean of $e_{t,OLS}$ is zero.

Properties of the estimators $\hat{\rho}_{OLS}(j)$ and $\hat{\rho}_{FD}(j)$ will depend on the assumptions on the process $v_t$ in model (1). In this paper, we only require $v_t$ to be stationary. This assumption allows us to assume $v_t$ to be a long memory $I(d)$ or ARFIMA($p,d,q$) process. Before presenting our main results, we first review some basic properties of the $I(d)$ process. A process $Y_t$ is said to be autoregressive fractionally integrated moving average process of order $p,d,q$, denoted as ARFIMA($p,d,q$) or $I(d)$, if it has a representation

$$\phi(L)(1 - L)^d (Y_t - \mu) = \theta(L)a_t,$$

where $\mu$ is the population mean, $L$ is the lag operator, $\phi(L)$ is a lag polynomial of degree $p$, $d$ is the differencing parameter which can be a fractional number and $\theta(L)$ is a lag polynomial of degree $q$. The innovation sequences $a_t$ is independent white noise with zero mean and variance $\sigma_a^2$.

**Assumption 2.1** The zeroes of $\phi(z)$ and $\theta(z)$ lie outside the unit circle and $\phi(z)$ and $\theta(z)$ have no common roots. Furthermore, $d \in (-0.5, 0.5)$.

The fractional differencing operator $(1 - L)^d$ has the following binomial expansion:

$$(1 - L)^d = \sum_{j=0}^{\infty} \psi_j L^j,$$

where $\psi_j = \Gamma(j - d)/\Gamma(j + 1)\Gamma(-d)$, and $\Gamma(\cdot)$ is the gamma function. The fractional white noise process is defined as

$$(1 - L)^d Y_t = a_t,$$

which is the simplest case of the ARFIMA model ($p = q = 0$). This process was first introduced by Granger (1980, 1981), Granger and Joyeux (1980), and Hosking (1981). We also refer to Baillie (1996) for a review on long memory and fractionally integrated processes.

To derive the asymptotic behaviors of $\hat{\rho}_{OLS}(j)$ and $\hat{\rho}_{FD}(j)$, we make the innovation sequences $a_t$ satisfy the following assumption throughout this paper. This assumption was first used by Davydov (1970).
Assumption 2.2 The white noise $a_t$ is independently and identically distributed with zero mean, and its moments satisfy the following condition: $E|a_t|^p < \infty$, with $p \geq \max\{4, -8d/(1 + 2d)\}$.

Without loss of generality, we also assume the initial value of the fractionally integrated processes, $v_0$, is zero. The independent and identical distribution assumption for $a_t$ is made to simplify our analysis and could be relaxed. In fact, given the preceding conditions on the stationary ARFIMA($p,d,q$) process, Hosking (1996, Theorem 8) showed that the exact order of magnitude of $\text{Var}(\sum_{t=1}^T \epsilon_t)$ is equal to $O(T^{1+2d})$. This asymptotic result is crucial to the derivation of the following theorem and it was established by Hosking (1996) assuming finite second moments of $a_t$.

Theorem 1 Under Assumptions 2.1 and 2.2 and $E(a_t^4) < \infty$, as $T \to \infty$, we have the following results:

$$\hat{\rho}_{\text{OLS}}(j) = \frac{\gamma_v(j) + O_p(T^{2d-1})}{\gamma_v(0) + O_p(T^{2d-1})} \xrightarrow{p} \rho_v(j) \quad \text{and} \quad \hat{\rho}_{\text{FD}}(j) = \frac{\gamma_v(j) + O_p(1)}{\gamma_v(0) + O_p(1)}.$$ 

Theorem 1 clearly indicates that $\hat{\rho}_{\text{FD}}(j)$ is not a consistent estimator for its population counterpart $\rho_v(j)$. Thus, it is not recommended to use $\hat{\rho}_{\text{FD}}(j)$ as a measure of autocorrelations for stationary detrended time series. Moreover, Theorem 1 shows that $\hat{\rho}_{\text{OLS}}(j)$ is a consistent estimator for $\rho_v(j)$, which justifies the use of OLS detrending in empirical applications.

3 Monte Carlo Experiment

We now conduct a Monte Carlo experiment to compare the finite sample performance of $\hat{\rho}_{\text{OLS}}(j)$ with that of $\hat{\rho}_{\text{FD}}(j)$ when three lag orders ($j = 3, 6, 9$) are used. We use a typical design of the Monte Carlo study. The experiment for each model is based on 1,000 replications with different sample sizes ($T$). We restrict the study to fractionally integrated I($d$) processes, or ARFIMA(0,$d$,0). To construct $T$ values of a stationary I($d$) process, we first generate $T$ independent values from the standard normal distribution and form a $T \times 1$ column vector $a = (a_1, \ldots, a_T)'$. We then calculate the $T$ analytic autocovariances of the I($d$) process, from which we construct the $T \times T$ variance-covariance matrix $\Sigma$ and compute its Cholesky decomposition, $\Sigma = CC'$. Finally, the vector $v = (v_1, \ldots, v_T)'$ of the $T$ realized values of the I($d$) process is defined by $v = Ca$. This algorithm was
Table 1: RMSE of autocorrelations for OLS detrending using alternative difference parameters $d$ and lags $j$.

<table>
<thead>
<tr>
<th>$T$</th>
<th>lag</th>
<th>I(-0.4)</th>
<th>I(-0.3)</th>
<th>I(-0.2)</th>
<th>I(-0.1)</th>
<th>I(0)</th>
<th>I(0.1)</th>
<th>I(0.2)</th>
<th>I(0.3)</th>
<th>I(0.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>3</td>
<td>0.1136</td>
<td>0.1111</td>
<td>0.1086</td>
<td>0.1066</td>
<td>0.1067</td>
<td>0.1156</td>
<td>0.1419</td>
<td>0.2261</td>
<td>0.4449</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.1143</td>
<td>0.1114</td>
<td>0.1084</td>
<td>0.1057</td>
<td>0.1043</td>
<td>0.1104</td>
<td>0.1331</td>
<td>0.2164</td>
<td>0.4392</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.1193</td>
<td>0.1163</td>
<td>0.1033</td>
<td>0.1104</td>
<td>0.1085</td>
<td>0.1131</td>
<td>0.1319</td>
<td>0.2113</td>
<td>0.4341</td>
</tr>
<tr>
<td>200</td>
<td>3</td>
<td>0.0775</td>
<td>0.0755</td>
<td>0.0995</td>
<td>0.0716</td>
<td>0.0712</td>
<td>0.0774</td>
<td>0.0955</td>
<td>0.1751</td>
<td>0.3905</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.0809</td>
<td>0.0789</td>
<td>0.0978</td>
<td>0.0750</td>
<td>0.0741</td>
<td>0.0789</td>
<td>0.0978</td>
<td>0.1715</td>
<td>0.3882</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.0817</td>
<td>0.0798</td>
<td>0.0971</td>
<td>0.0763</td>
<td>0.0755</td>
<td>0.0799</td>
<td>0.0971</td>
<td>0.1690</td>
<td>0.3860</td>
</tr>
<tr>
<td>300</td>
<td>3</td>
<td>0.0623</td>
<td>0.0607</td>
<td>0.0591</td>
<td>0.0577</td>
<td>0.0575</td>
<td>0.0627</td>
<td>0.0817</td>
<td>0.1500</td>
<td>0.3592</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.0664</td>
<td>0.0647</td>
<td>0.0630</td>
<td>0.0614</td>
<td>0.0606</td>
<td>0.0642</td>
<td>0.0801</td>
<td>0.1466</td>
<td>0.3568</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.0621</td>
<td>0.0613</td>
<td>0.0599</td>
<td>0.0587</td>
<td>0.0581</td>
<td>0.0615</td>
<td>0.0762</td>
<td>0.1419</td>
<td>0.3538</td>
</tr>
<tr>
<td>400</td>
<td>3</td>
<td>0.0540</td>
<td>0.0526</td>
<td>0.0513</td>
<td>0.0501</td>
<td>0.0500</td>
<td>0.0547</td>
<td>0.0722</td>
<td>0.1363</td>
<td>0.3406</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.0577</td>
<td>0.0563</td>
<td>0.0548</td>
<td>0.0535</td>
<td>0.0527</td>
<td>0.0560</td>
<td>0.0706</td>
<td>0.1330</td>
<td>0.3385</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.0548</td>
<td>0.0534</td>
<td>0.0520</td>
<td>0.0508</td>
<td>0.0502</td>
<td>0.0535</td>
<td>0.0677</td>
<td>0.1302</td>
<td>0.3369</td>
</tr>
<tr>
<td>500</td>
<td>3</td>
<td>0.0493</td>
<td>0.0481</td>
<td>0.0469</td>
<td>0.0459</td>
<td>0.0460</td>
<td>0.0506</td>
<td>0.0669</td>
<td>0.1273</td>
<td>0.3271</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.0506</td>
<td>0.0493</td>
<td>0.0480</td>
<td>0.0468</td>
<td>0.0463</td>
<td>0.0494</td>
<td>0.0634</td>
<td>0.1229</td>
<td>0.3244</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.0480</td>
<td>0.0468</td>
<td>0.0455</td>
<td>0.0444</td>
<td>0.0439</td>
<td>0.0468</td>
<td>0.0603</td>
<td>0.1201</td>
<td>0.3229</td>
</tr>
</tbody>
</table>

suggests by McLeod and Hipel (1978) and Hosking (1984). Furthermore, 50 additional values are generated in order to obtain random starting values. Finally, we choose $\alpha = 1$ and $\beta = 0.03$ to generate the series $y_t$ using model (1).

Tables 1 and 2 contain the simulation results for the root mean squared error (RMSE) of $\hat{\rho}_{FD}(j)$ and $\hat{\rho}_{OLS}(j)$. We can draw some conclusions from them. First, the RMSE of $\hat{\rho}_{OLS}(j)$ decreases monotonically with the increase of the sample size. This phenomenon coincides with the result in Theorem 1 that $\hat{\rho}_{OLS}(j)$ is a consistent estimator of $\gamma_v(j)$. Second, no matter which lag number $j$ and differencing parameter $d$ is used in the experiment, the RMSE of $\hat{\rho}_{FD}(j)$ does not diminish monotonically. We should emphasize that this holds also for the short memory case, I(0). These results confirm the theoretical result of Theorem 1.

To illustrate our simulation results graphically, Figure 1 and Figure 2 show the results in Tables 1 and 2 for the case $d = 0.3$. Five additional sample sizes are added in the simulations, i.e., $T = 600, 700, 800, 900, 1000$. Again, the performance of the estimators
Table 2: RMSE of autocorrelations for FD detrending using alternative difference parameters $d$ and lags $j$.

<table>
<thead>
<tr>
<th>T</th>
<th>lag</th>
<th>I(-0.4)</th>
<th>I(-0.3)</th>
<th>I(-0.2)</th>
<th>I(-0.1)</th>
<th>I(0)</th>
<th>I(0.1)</th>
<th>I(0.2)</th>
<th>I(0.3)</th>
<th>I(0.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>3</td>
<td>0.2922</td>
<td>0.2911</td>
<td>0.2896</td>
<td>0.2870</td>
<td>0.2778</td>
<td>0.2536</td>
<td>0.2405</td>
<td>0.3775</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.2742</td>
<td>0.2729</td>
<td>0.2714</td>
<td>0.2690</td>
<td>0.2645</td>
<td>0.2606</td>
<td>0.2381</td>
<td>0.2277</td>
<td>0.3728</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.2562</td>
<td>0.2551</td>
<td>0.2538</td>
<td>0.2520</td>
<td>0.2484</td>
<td>0.2454</td>
<td>0.2253</td>
<td>0.2187</td>
<td>0.3706</td>
</tr>
<tr>
<td>200</td>
<td>3</td>
<td>0.2902</td>
<td>0.2914</td>
<td>0.2695</td>
<td>0.2933</td>
<td>0.2923</td>
<td>0.2925</td>
<td>0.2695</td>
<td>0.2394</td>
<td>0.3285</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.2763</td>
<td>0.2782</td>
<td>0.2605</td>
<td>0.2816</td>
<td>0.2814</td>
<td>0.2822</td>
<td>0.2605</td>
<td>0.2318</td>
<td>0.3253</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.2666</td>
<td>0.2682</td>
<td>0.2504</td>
<td>0.2709</td>
<td>0.2707</td>
<td>0.2713</td>
<td>0.2504</td>
<td>0.2228</td>
<td>0.3217</td>
</tr>
<tr>
<td>300</td>
<td>3</td>
<td>0.2981</td>
<td>0.2968</td>
<td>0.2958</td>
<td>0.2948</td>
<td>0.2930</td>
<td>0.2941</td>
<td>0.2735</td>
<td>0.2421</td>
<td>0.3020</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.2910</td>
<td>0.2898</td>
<td>0.2888</td>
<td>0.2880</td>
<td>0.2865</td>
<td>0.2877</td>
<td>0.2680</td>
<td>0.2371</td>
<td>0.2991</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.2840</td>
<td>0.2830</td>
<td>0.2823</td>
<td>0.2818</td>
<td>0.2809</td>
<td>0.2825</td>
<td>0.2635</td>
<td>0.2327</td>
<td>0.2959</td>
</tr>
<tr>
<td>400</td>
<td>3</td>
<td>0.2874</td>
<td>0.2849</td>
<td>0.2826</td>
<td>0.2802</td>
<td>0.2774</td>
<td>0.2780</td>
<td>0.2590</td>
<td>0.2289</td>
<td>0.2825</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.2845</td>
<td>0.2823</td>
<td>0.2801</td>
<td>0.2780</td>
<td>0.2755</td>
<td>0.2762</td>
<td>0.2578</td>
<td>0.2280</td>
<td>0.2813</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.2801</td>
<td>0.2778</td>
<td>0.2757</td>
<td>0.2737</td>
<td>0.2714</td>
<td>0.2724</td>
<td>0.2546</td>
<td>0.2252</td>
<td>0.2789</td>
</tr>
<tr>
<td>500</td>
<td>3</td>
<td>0.2994</td>
<td>0.3026</td>
<td>0.3061</td>
<td>0.3097</td>
<td>0.3123</td>
<td>0.3179</td>
<td>0.2992</td>
<td>0.2609</td>
<td>0.2749</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.2961</td>
<td>0.2992</td>
<td>0.3026</td>
<td>0.3062</td>
<td>0.3088</td>
<td>0.3141</td>
<td>0.2957</td>
<td>0.2573</td>
<td>0.2720</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.2928</td>
<td>0.2960</td>
<td>0.2996</td>
<td>0.3033</td>
<td>0.3061</td>
<td>0.3115</td>
<td>0.2936</td>
<td>0.2556</td>
<td>0.2708</td>
</tr>
</tbody>
</table>

For finite samples corresponds to the asymptotic result.

From Tables 1 and 2, and Figures 1 and 2, it is clear that the FD residuals should not be used to construct autocorrelation functions. In other words, OLS detrending is the only consistent estimator to be considered. Therefore, we now analyze more in detail the finite sample performance of $\hat{\rho}_{OLS}(j)$. Theorem 1 shows that the bias of $\hat{\rho}_{OLS}(j)$ depends on the magnitude of $d$. However, we do not know the actual size of the bias. To clarify the usefulness of OLS residuals in calculating autocorrelation functions, we compare the RMSE of $\hat{\rho}_{OLS}(j)$ with $\hat{\rho}_{v}(j)$ which is constructed by using the true error term $v_t$. The experiment is performed with a sample size of $T = 200$ and the results are shown in Figures 3 and 4 for lags three and nine, respectively. These figures clearly indicate that $\hat{\rho}_{v}(j)$ and $\hat{\rho}_{OLS}(j)$ are close to each other independent of the lag order. Therefore, the use of OLS detrending in empirical analysis is justified both by asymptotic theory and by our simulation results.
4 Empirical Example

We analyze daily observations of volume in Dow Jones stocks from January 2, 1973 to May 13, 2003, a total of 7665 observations. We consider dollar volume, defined by the stock price multiplied by the number of stocks traded. Prediction of volume is important for liquidity assessments and optimal trading strategies. Model-based predictions will mainly use the autocorrelation structure of the data, so that it is important to have reliable estimates of the ACF of volume. Moreover, volume is closely linked to volatility, often explained by a common subordinated information arrival process, as suggested e.g. by Tauchen and Pitts (1983). Thus, understanding the dynamics of volume, which is observed, may help to better understand the dynamics of unobserved volatility. See e.g. Gallant, Rossi an Tauchen (1992) for a broad overview of the importance of volume for the dynamics of stock prices.

To pick two typical examples, the logarithm of volume is plotted in Figures 5 and 7 for McDonald’s and Philip Morris, respectively. Clearly, there is an upward trend in both series and it may be approximated by a linear trend. These are typical examples in the sense that the difference of the estimates of $\beta$ using the two methods, reported in Table 3, corresponds roughly to the average difference over the series. Figures 6 and 8 depict the ACF of residuals for both OLS and FD detrending. We see a clear difference between both ACF, where the ACF of OLS residuals decays faster than that of FD residuals. That is, the persistence is overestimated using the FD method. In general, we can identify differences of both ACF whenever there are differences in the estimates of $\beta$. For the case of IBM with only small differences in estimated $\beta$ coefficients, also the differences of ACF is small. On the other extreme is Homedepot with a huge difference in $\beta$ estimates that translates into a very slow decay of the ACF using FD but a fast one using OLS.

We also report the estimates of the ARFIMA($p, d, q$) models in Table 3, where the order of $p$ and $q$ was either one or zero and chosen by the AIC criterion. In all cases, the estimate of $d$ is significantly larger zero, in all cases between 0.34 and 0.49, indicating a strong degree of long memory but stationarity. The conclusion is that one should be careful in using the FD detrending method when using it for prediction of volume in financial markets, because autocorrelation estimates may not be accurate.
5 Conclusion

This paper considers the impact of detrending methods on autocorrelation estimation. The ordinary least squares (OLS) detrending and the first difference (FD) detrending methods are considered. Our analysis reveals that \( \hat{\rho}_{\text{OLS}}(j) \xrightarrow{p} \rho_v(j) \) as long as \( v_t \) is stationary. However, the FD detrending results in inconsistent autocorrelation functions. Moreover, our simulation results show that the RMSE of \( \hat{\rho}_{\text{OLS}}(j) \) decreases monotonically with the increase of sample size. Furthermore, no matter what lag number and differencing parameter is used in the experiment, the RMSE of \( \hat{\rho}_{\text{FD}}(j) \) does not diminish monotonically. Therefore, the predictions made in Theorem 1 that the FD detrending will result in inconsistent autocorrelation functions is clearly supported in our simulations. This suggests that the FD detrending should be employed more carefully when there is evidence for trend stationarity.

Appendix

Proof of Theorem 1

Let us first consider the asymptotic properties of \( \hat{\rho}_{\text{OLS}}(j) \). We note that

\[
\hat{\beta}_{\text{OLS}} - \beta = \frac{\sum_{t=1}^{T} (v_t - \bar{v})(t - \bar{t})}{\sum_{t=1}^{T} (t - \bar{t})^2} = \frac{\sum_{t=1}^{T} tv_t - \frac{1}{T} \left( \sum_{t=1}^{T} t \right) \left( \sum_{t=1}^{T} v_t \right)}{\sum_{t=1}^{T} (t - \bar{t})^2} = O_p\left(T^{d-1.5}\right),
\]

by using items 4 and 10 of Lemma 1 of Tsay and Chung (2000). We can rewrite \( \hat{\rho}_{\text{OLS}}(j) \) as:

\[
\hat{\rho}_{\text{OLS}}(j) = \frac{1}{T} \sum_{t=j+1}^{T} (e_{t,\text{OLS}})(e_{t-j,\text{OLS}}) \cdot \frac{1}{T} \sum_{t=1}^{T} e_{t,\text{OLS}}^2.
\]
For the denominator of \( \hat{\rho}_{OLS}(j) \), we have
\[
\frac{1}{T} \sum_{t=1}^{T} e_{t,OLS}^2 = \frac{1}{T} \sum_{t=1}^{T} (y_t - \hat{\alpha}_{OLS} - \hat{\beta}_{OLS} t)^2 \\
= \frac{1}{T} \sum_{t=1}^{T} (v_t - \overline{v})^2 - \frac{1}{T} \left( \hat{\beta}_{OLS} - \beta \right) \sum_{t=1}^{T} (t - \overline{t})^2 \\
= \frac{1}{T} \sum_{t=1}^{T} (v_t - \overline{v})^2 + O_p(T^{2d-1}) \\
= \gamma_v(0) + O_p(T^{2d-1}) \xrightarrow{p} \gamma_v(0)
\]
by Lemma 1, item 6 of Tsay and Chung (2000) and the preceding results.

For the numerator of \( \hat{\rho}_{OLS}(j) \), we have
\[
\sum_{t=j+1}^{T} (e_{t,OLS})(e_{t-j,OLS}) = \sum_{t=j+1}^{T} (v_t - \overline{v})(v_{t-j} - \overline{v}) \\
- \left( \hat{\beta}_{OLS} - \beta \right) \sum_{t=j+1}^{T} (v_t - \overline{v})(t - j - \overline{t}) \\
- \left( \hat{\beta}_{OLS} - \beta \right) \sum_{t=j+1}^{T} (v_{t-j} - \overline{v})(t - \overline{t}) \\
+ \left( \hat{\beta}_{OLS} - \beta \right)^2 \sum_{t=j+1}^{T} (t - \overline{t})(t - j - \overline{t}) \\
\equiv A + B + C + D.
\]

For the term \( A \), we note that
\[
\frac{1}{T} \sum_{t=j+1}^{T} (v_t - \overline{v})(v_{t-j} - \overline{v}) = \gamma_v(j) + O_p(T^{2d-1}) \xrightarrow{p} \gamma_v(j), \tag{2}
\]
by Theorem 3 of Hosking (1996).

For the term \( B \), first we have
\[
\sum_{t=j+1}^{T} (v_t - \overline{v})(t - j - \overline{t}) = \sum_{t=j+1}^{T} (v_t - \overline{v})(t - \overline{t}) - \sum_{t=j+1}^{T} j(v_t - \overline{v}) = O_p(T^{d+1.5}), \tag{3}
\]
because \( j/T \to 0 \) as \( T \to \infty \). This implies that \( B = O_p(T^{d-1.5})O_p(T^{d+1.5}) = O_p(T^{2d}) \).
For the term $C$, similarly, we have
\[
\sum_{t=j+1}^{T} (v_{t-j} - \bar{v}) (t - \bar{t}) = \sum_{t=j+1}^{T} tv_{t-j} - \sum_{t=j+1}^{T} t\bar{v} - \sum_{t=j+1}^{T} v_{t-j}\bar{t} + \sum_{t=j+1}^{T} t\bar{v} = \sum_{t=j+1}^{T} tv_{t-j} + O_p(T^{d+1.5}) + O_p(T^{d+1.5}) + O_p(T^{d+1.5}).
\]

Furthermore, we also note that
\[
\sum_{t=j+1}^{T} tv_{t-j} = \sum_{t=1}^{T-j} tv_t + j \sum_{t=1}^{T-j} v_t = O_p(T^{d+1.5}),
\]
and we have
\[
\sum_{t=j+1}^{T} (v_{t-j} - \bar{v}) (t - \bar{t}) = O_p(T^{d+1.5}).
\]

Therefore, we prove that $C = O_p(T^{d-1.5}) O_p(T^{d+1.5}) = O_p(T^{2d})$.

For the term $D$, we note that
\[
\sum_{t=j+1}^{T} (t - \bar{t}) (t - j - \bar{t}) = \sum_{t=j+1}^{T} (t - \bar{t}) (t - \bar{t}) - j \sum_{t=j+1}^{T} (t - \bar{t})^2 = O(T^3),
\]
and we prove that $D = O(T^{2d-3}) O_p(T^3) = O_p(T^{2d})$. Combining the asymptotic properties of $A$, $B$, $C$, and $D$, we prove that
\[
\frac{1}{T} \sum_{t=j+1}^{T} (e_t,OLS)(e_{t-j},OLS) = \frac{1}{T} \sum_{t=j+1}^{T} (v_t - \bar{v})(v_{t-j} - \bar{v}) + O_p(T^{2d-1})
\]
\[
= \gamma_v(j) + O_p(T^{2d-1}) \overset{p}{\rightarrow} \gamma_v(j).
\]

Given the asymptotic properties of the numerator and the denominator of $\hat{\rho}_{OLS}(j)$, we have using Slutsky’s theorem that
\[
\hat{\rho}_{OLS}(j) = \frac{\gamma_v(j) + O_p(T^{2d-1})}{\gamma_v(0) + O_p(T^{2d-1})} \overset{p}{\rightarrow} \rho_v(j).
\]

We now consider the asymptotic properties of $\hat{\rho}_{FD}(j)$. Note that
\[
\Delta y_t = \beta + \Delta v_t.
\]

For the FD estimator, we have
\[
\hat{\beta}_{FD} - \beta = \frac{1}{T-1} \sum_{t=2}^{T} \Delta v_t = O_p(T^{-1}).
\]
by Theorem 3 of Tsay (2000). Moreover, rewrite $\hat{\rho}_{FD}(j)$ as

$$\hat{\rho}_{FD}(j) = \frac{1}{T} \sum_{t=j+1}^{T} (e_{t,FD} - \bar{\epsilon}_{FD})(e_{t-j,FD} - \bar{\epsilon}_{FD})$$

For the denominator of $\hat{\rho}_{FD}(j)$, we have

$$\frac{1}{T} \sum_{t=1}^{T} (e_{t,FD} - \bar{\epsilon}_{FD})^2 = \frac{1}{T} \sum_{t=1}^{T} (y_t - \hat{\beta}_{FD}t - \bar{y} + \hat{\beta}_{FD}\bar{t})^2$$

$$= \frac{1}{T} \sum_{t=1}^{T} (v_t - \bar{v})^2$$

$$- \frac{2}{T} (\hat{\beta}_{FD} - \beta) \sum_{t=1}^{T} (v_t - \bar{v})(t - \bar{t}) + \frac{1}{T} (\hat{\beta}_{FD} - \beta)^2 \sum_{t=1}^{T} (t - \bar{t})^2$$

$$= \frac{1}{T} \sum_{t=1}^{T} (v_t - \bar{v})^2 + O_p(T^{d-0.5}) + O_p(1).$$

For the numerator of $\hat{\rho}_{FD}(j)$, we have

$$\sum_{t=j+1}^{T} (e_{t,FD} - \bar{\epsilon}_{FD})(e_{t-j,FD} - \bar{\epsilon}_{FD}) = \sum_{t=j+1}^{T} (v_t - \bar{v})(v_{t-j} - \bar{v})$$

$$- (\hat{\beta}_{FD} - \beta) \sum_{t=j+1}^{T} (v_t - \bar{v})(t - j - \bar{t})$$

$$- (\hat{\beta}_{FD} - \beta) \sum_{t=j+1}^{T} (v_{t-j} - \bar{v})(t - \bar{t})$$

$$+ (\hat{\beta}_{FD} - \beta)^2 \sum_{t=j+1}^{T} (t - \bar{t})(t - j - \bar{t})$$

$$\equiv A^* + B^* + C^* + D^*.$$

For the term $A^*$, we note that $A^* \equiv A$, and we prove that

$$\frac{1}{T} \sum_{t=j+1}^{T} (v_t - \bar{v})(v_{t-j} - \bar{v}) = \gamma_v(j) + O_p(T^{2d-1}) \xrightarrow{p} \gamma_v(j).$$

For the term $B^*$, we note that

$$\sum_{t=j+1}^{T} (v_t - \bar{v})(t - j - \bar{t}) = O_p(T^{d+1.5}),$$

12
by Equation (3). This implies that $B^* = O_p(T^{-1}) O_p(T^{d+1.5}) = O_p(T^{d+0.5})$.

For the term $C^*$, we note that
\[\sum_{t=j+1}^{T} (v_{t-j} - \bar{v})(t - \bar{t}) = O_p(T^{d+1.5}) ,\]
by Equation (5), and $C^* = O_p(T^{-1}) O_p(T^{1.5+d}) = O_p(T^{d+0.5})$.

For the term $D^*$, we note that
\[\sum_{t=j+1}^{T} (t - \bar{t})(t - j - \bar{t}) = O(T^3) ,\]
by Equation (6). Therefore, $D^* = O_p(T^{-2}) O(T^3) = O_p(T)$. Combining the asymptotic properties of $A^*$, $B^*$, $C^*$, and $D^*$, we have
\[\frac{1}{T} \sum_{t=j+1}^{T} (e_{t,FD} - \bar{v})(e_{t-j,FD} - \bar{v}) = \gamma_v(j) + O_p(1) .\]

Finally, given the asymptotic properties of the numerator and the denominator of $\hat{\rho}_{FD}(j)$, we have shown that
\[\hat{\rho}_{FD}(j) = \frac{\gamma_v(j) + O_p(1)}{\gamma_v(0) + O_p(1)} .\] (8)

Therefore, the FD detrending will result in inconsistent estimates of the autocorrelation functions, as stated.

References


<table>
<thead>
<tr>
<th>Company</th>
<th>d</th>
<th>AR</th>
<th>MA</th>
<th>$\beta_{OLS}$</th>
<th>$\beta_{FD}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP</td>
<td>0.4236(0.012)</td>
<td>0.0141(0.018)</td>
<td>0.0013</td>
<td>0.0014</td>
<td></td>
</tr>
<tr>
<td>AMEX</td>
<td>0.4130(0.0128)</td>
<td>-0.0618(0.0192)</td>
<td>0.0009</td>
<td>0.0006</td>
<td></td>
</tr>
<tr>
<td>MCDONALDS</td>
<td>0.3936(0.011)</td>
<td></td>
<td></td>
<td>0.0013</td>
<td>0.0010</td>
</tr>
<tr>
<td>PHMORRIS</td>
<td>0.4272(0.0122)</td>
<td></td>
<td></td>
<td>0.0016</td>
<td>0.0013</td>
</tr>
<tr>
<td>UTECH</td>
<td>0.4347(0.012)</td>
<td>-0.1094(0.018)</td>
<td>0.0012</td>
<td>0.0013</td>
<td></td>
</tr>
<tr>
<td>JOHNSON</td>
<td>0.3902(0.0079)</td>
<td></td>
<td></td>
<td>0.0014</td>
<td>0.0012</td>
</tr>
<tr>
<td>ALCOA</td>
<td>0.3789(0.0125)</td>
<td>-0.04(0.017)</td>
<td>0.0011</td>
<td>0.0012</td>
<td></td>
</tr>
<tr>
<td>SBC</td>
<td>0.3712(0.0116)</td>
<td></td>
<td></td>
<td>0.0015</td>
<td>0.0014</td>
</tr>
<tr>
<td>CATERPILLAR</td>
<td>0.3858(0.008)</td>
<td></td>
<td></td>
<td>0.0008</td>
<td>0.0009</td>
</tr>
<tr>
<td>MICROSOFT</td>
<td>0.4858(0.0099)</td>
<td></td>
<td></td>
<td>0.0028</td>
<td>0.0019</td>
</tr>
<tr>
<td>MERCK</td>
<td>0.4181(0.0116)</td>
<td></td>
<td></td>
<td>0.0015</td>
<td>0.0012</td>
</tr>
<tr>
<td>DUPONT</td>
<td>0.3927(0.0120)</td>
<td></td>
<td></td>
<td>0.0012</td>
<td>0.0011</td>
</tr>
<tr>
<td>HOMEDEPOT</td>
<td>0.3927(0.0120)</td>
<td></td>
<td></td>
<td>0.0023</td>
<td>0.0001</td>
</tr>
<tr>
<td>IBM</td>
<td>0.4747(0.0099)</td>
<td></td>
<td></td>
<td>0.0009</td>
<td>0.0009</td>
</tr>
<tr>
<td>HONEYWELL</td>
<td>0.4147(0.0112)</td>
<td>-0.037(0.01573)</td>
<td>0.0011</td>
<td>0.0010</td>
<td></td>
</tr>
<tr>
<td>PROCTERGAMBLE</td>
<td>0.3836(0.0115)</td>
<td>-0.0507(0.0156)</td>
<td>0.0012</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td>WALMART</td>
<td>0.4715(0.0135)</td>
<td>-0.247(0.1074)</td>
<td>0.0024</td>
<td>0.0021</td>
<td></td>
</tr>
<tr>
<td>INTEL</td>
<td>0.4612(0.0127)</td>
<td>-0.1375(0.0193)</td>
<td>0.0022</td>
<td>0.0018</td>
<td></td>
</tr>
<tr>
<td>GENELECTRIC</td>
<td>0.3836(0.1115)</td>
<td>-0.0512(0.0156)</td>
<td>0.0014</td>
<td>0.0012</td>
<td></td>
</tr>
<tr>
<td>CITI</td>
<td>0.4636(0.012)</td>
<td>-0.1355(0.0193)</td>
<td>0.0018</td>
<td>0.0015</td>
<td></td>
</tr>
<tr>
<td>EXXON</td>
<td>0.412(0.0133)</td>
<td>-0.03(0.015)</td>
<td>0.0010</td>
<td>0.0010</td>
<td></td>
</tr>
<tr>
<td>MMM</td>
<td>0.3447(0.008)</td>
<td></td>
<td></td>
<td>0.0010</td>
<td>0.0010</td>
</tr>
<tr>
<td>JPMORGAN</td>
<td>0.3811(0.010)</td>
<td></td>
<td></td>
<td>0.0016</td>
<td>0.0015</td>
</tr>
<tr>
<td>BOEING</td>
<td>0.4551(0.01093)</td>
<td>-0.04(0.017)</td>
<td>0.0014</td>
<td>0.0010</td>
<td></td>
</tr>
<tr>
<td>DISNEY</td>
<td>0.4643(0.0123)</td>
<td>-0.0390(0.0172)</td>
<td>0.0014</td>
<td>0.0010</td>
<td></td>
</tr>
<tr>
<td>EKODAK</td>
<td>0.3715(0.1179)</td>
<td>0.0948(0.0167)</td>
<td>0.0007</td>
<td>0.0006</td>
<td></td>
</tr>
<tr>
<td>AT&amp;T</td>
<td>0.4563(0.0113)</td>
<td></td>
<td></td>
<td>0.0006</td>
<td>0.0002</td>
</tr>
<tr>
<td>COCACOLA</td>
<td>0.4109(0.0125)</td>
<td>-0.0327(0.0147)</td>
<td>0.0015</td>
<td>0.0013</td>
<td></td>
</tr>
<tr>
<td>INTLPAPER</td>
<td>0.3710(0.0127)</td>
<td></td>
<td></td>
<td>0.0010</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

Table 3: Parameter estimates of ARFIMA models for the Dow Jones volume data. d is the fractional difference parameter, AR and MA the autoregressive and moving average component, respectively, and $\beta_{OLS}$ and $\beta_{FD}$ are the estimates of the trend coefficient using OLS and FD detrending. Standard errors are in parentheses.
Figure 1: Root mean squared error of ACF estimates using OLS detrending.

Figure 2: Root mean squared error of ACF estimates using FD detrending.
Figure 3: RMSE of autocorrelations estimated using OLS detrending (solid line) and the true simulated values of $v_t$ (dashed line) for a sample size of $T = 200$ and lag 3.

Figure 4: RMSE of autocorrelations estimated using OLS detrending (solid line) and the true simulated values of $v_t$ (dashed line) for a sample size of $T = 200$ and lag 9.
Figure 5: Time series of McDonalds daily log turnover, 1973 to 2003.

Figure 6: ACF of log turnover of McDonalds using OLS (solid) and FD (dashed) detrending.
Figure 7: Time series of Philip Morris daily log turnover, 1973 to 2003.

Figure 8: ACF of log turnover of Philip Morris using OLS (solid) and FD (dashed) detrending.
Recent titles

CORE Discussion Papers

2008/38. Helmhut CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Habit formation and labor supply.
2008/42. Raouf BOUCEKKINE, Rodolphe DESBORDES and Hélène LATZER. How do epidemics induce behavioral changes?
2008/43. David DE LA CROIX and Marie VANDER DONCKT. Would empowering women initiate the demographic transition in least-developed countries?
2008/44. Geoffrey CARUSO, Dominique PEETERS, Jean CAVAILHES and Mark ROUNSEVELL. Space-time patterns of urban sprawl, a 1D cellular automata and microeconomic approach.
2008/45. Taoufik BOUEZMARNI, Jeroen V. K. ROMBOUTS and Abderrahim TAAMOUTI. Asymptotic properties of the Bernstein density copula for dependent data.
2008/46. Joe THARAKAN and Jean-Philippe TROPEANO. On the impact of labor market matching on regional disparities.
2008/47. Shin-Huei WANG and Cheng HSIAO. An easy test for two stationary long processes being uncorrelated via AR approximations.
2008/54. Sergei PEKARSKI. Budget deficits and inflation feedback. 2008/55. Raouf BOUCEKKINE, Jacek B. KRAWCZYK and Thomas VALLEE. Towards an understanding of tradeoffs between regional wealth, tightness of a common environmental constraint and the sharing rules.
2008/58. Marie-Louise LEROUX, Pierre PESTIEAU and Gregory PONTHIERE. Should we subsidize longevity?
2008/59. J. Roderick McCORIE. The role of Skorokhod space in the development of the econometric analysis of time series.
2008/60. Yu. NESTEROV. Barrier subgradient method.
2008/61. Thierry BRECHET, Johan EVCKMANS, François GERARD, Philippe MARBAIX, Henry TULKENS and Jean-Pascal VAN YPERSELE. The impact of the unilateral EU commitment on the stability of international climate agreements.
2008/62. Giorgia OGGIONI and Yves SMEERS. Average power contracts can mitigate carbon leakage.
Recent titles

CORE Discussion Papers - continued

2008/64. Nicolas GILLIS and François GLINEUR. Nonnegative factorization and the maximum edge biclique problem.
2008/70. Michel JOURNÉE, Yuriii NESTEROV, Peter RICHTÁRIK and Rodolphe SEPULCHRE. Generalized power method for sparse principal component analysis.
2008/71. Toshihiro OKUBO and Pierre M. PICARD. Firms' location under taste and demand heterogeneity.
2008/72. Iwan MEIER and Jeroen V.K. ROMBOUTS. Style rotation and performance persistence of mutual funds.

Books

V. GINSBURGH and D. THROSBY (eds.) (2006), Handbook of the economics of art and culture. Amsterdam, Elsevier.

CORE Lecture Series

R. AMIR (2002), Supermodularity and complementarity in economics.
R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.