2008/73

■

Estimating autocorrelations in the presence of deterministic trends

Shin-Huei Wang and Christian M. Hafner

CORE DISCUSSION PAPER
2008/73

Estimating autocorrelations in the presence of deterministic trends

Shin-Huei WANG [1] and Christian M. HAFNER [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

---

[1] Université catholique de Louvain, CORE, B-1348 Louvain-la-Neuve, Belgium. E-mail: shin-huei.wang@uclouvain.be

[2] Université catholique de Louvain, Institut de Statistique and CORE, B-1348 Louvain-la-Neuve, Belgium. E-mail: Christian.hafner@uclouvain.be

1 Introduction

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

$$y_t = \alpha + \beta t + v_t, \quad t = 1, \ldots, T \tag{1}$$

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 β. In this paper, we turn our attention to the autocorrelations of v_t, defined as: $\widehat{\rho}_v(j) = \sum_{t=j+1}^{T}(v_t - \overline{v})(v_{t-j} - \overline{v})/\sum_{t=1}^{T}(v_t - \overline{v})^2$, 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, $\widehat{\rho}_y(j) = \sum_{t=j+1}^{T}(y_t - \overline{y})(y_{t-j} - \overline{y})/\sum_{t=1}^{T}(y_t - \overline{y})^2$, then the information contained in $\widehat{\rho}_y(j)$ will not be the one we need. In fact, we can prove that $\widehat{\rho}_y(j) \xrightarrow{p} 1$ 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 β 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, $\widehat{\rho}_{OLS}(j)$ and $\widehat{\rho}_{FD}(j)$, respectively. We would like to show that $\widehat{\rho}_{OLS}(j)$ and $\widehat{\rho}_{FD}(j)$ 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 $\widehat{\rho}_{OLS}(j) \xrightarrow{p} \rho_v(j)$ as long as v_t is stationary. On the other hand, the consistency of $\widehat{\rho}_{FD}(j)$ 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 $\widehat{\rho}_{OLS}(j)$ decreases monotonically with the sample size. Furthermore, no matter what lag number and differencing parameter is used in the experiment, the RMSE of $\widehat{\rho}_{FD}(j)$ 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 $\widehat{\rho}_{OLS}(j)$ and $\widehat{\rho}_{FD}(j)$ 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 $\widehat{\alpha}_{OLS}$ and $\widehat{\beta}_{OLS}$ to obtain OLS residuals $e_{t,OLS} = y_t - \widehat{\alpha}_{OLS} - \widehat{\beta}_{OLS}t$. FD detrending uses the estimator $\widehat{\beta}_{FD} = \frac{1}{T-1}\sum_{t=2}^{T}\Delta y_t$ to obtain residuals $e_{t,FD} = y_t - \widehat{\beta}_{FD}t$.

From these residuals, we construct two autocorrelation functions,

$$\widehat{\rho}_{OLS}(j) = \frac{\sum\limits_{t=j+1}^{T} e_{t,OLS} e_{t-j,OLS}}{\sum\limits_{t=1}^{T} e_{t,OLS}^2}, \quad \widehat{\rho}_{FD}(j) = \frac{\sum\limits_{t=j+1}^{T} \left(e_{t,FD} - \overline{e}_{FD}\right)\left(e_{t-j,FD} - \overline{e}_{FD}\right)}{\sum\limits_{t=1}^{T} \left(e_{t,FD} - \overline{e}_{FD}\right)^2},$$

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

Properties of the estimators $\widehat{\rho}_{OLS}(j)$ and $\widehat{\rho}_{FD}(j)$ 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 μ is the population mean, L is the lag operator, $\phi(L)$ is a lag polynomial of degreee 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 σ_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 $\widehat{\rho}_{OLS}(j)$ and $\widehat{\rho}_{FD}(j)$, 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:*

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

Theorem 1 clearly indicates that $\widehat{\rho}_{FD}(j)$ is not a consistent estimator for its population counterpart $\rho_v(j)$. Thus, it is not recommended to use $\widehat{\rho}_{FD}(j)$ as a measure of autocorrelations for stationary detrended time series. Moreover, Theorem 1 shows that $\widehat{\rho}_{OLS}(j)$ 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 $\widehat{\rho}_{OLS}(j)$ with that of $\widehat{\rho}_{FD}(j)$ 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 Σ 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

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

Table 1: RMSE of autocorrelations for OLS detrending using alternative difference parameters d and lags j.

suggested 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 $\widehat{\rho}_{FD}(j)$ and $\widehat{\rho}_{OLS}(j)$. We can draw some conclusions from them. First, the RMSE of $\widehat{\rho}_{OLS}(j)$ decreases monotonically with the increase of the sample size. This phenomenon coincides with the result in Theorem 1 that $\widehat{\rho}_{OLS}(j)$ 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 $\widehat{\rho}_{FD}(j)$ 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

T	lag	I(-0.4)	I(-0.3)	I(-0.2)	I(-0.1)	I(0)	I(0.1)	I(0.2)	I(0.3)	I(0.4)
100	3	0.2922	0.2911	0.2896	0.2870	0.2821	0.2778	0.2536	0.2405	0.3775
	6	0.2742	0.2729	0.2714	0.2690	0.2645	0.2606	0.2381	0.2277	0.3728
	9	0.2562	0.2551	0.2538	0.2520	0.2484	0.2454	0.2253	0.2187	0.3706
200	3	0.2902	0.2914	0.2695	0.2933	0.2923	0.2925	0.2695	0.2394	0.3285
	6	0.2763	0.2782	0.2605	0.2816	0.2814	0.2822	0.2605	0.2318	0.3253
	9	0.2666	0.2682	0.2504	0.2709	0.2707	0.2713	0.2504	0.2228	0.3217
300	3	0.2981	0.2968	0.2958	0.2948	0.2930	0.2941	0.2735	0.2421	0.3020
	6	0.2910	0.2898	0.2888	0.2880	0.2865	0.2877	0.2680	0.2371	0.2991
	9	0.2840	0.2830	0.2823	0.2818	0.2809	0.2825	0.2635	0.2327	0.2959
400	3	0.2874	0.2849	0.2826	0.2802	0.2774	0.2780	0.2590	0.2289	0.2825
	6	0.2845	0.2823	0.2801	0.2780	0.2755	0.2762	0.2578	0.2280	0.2813
	9	0.2801	0.2778	0.2757	0.2737	0.2714	0.2724	0.2546	0.2252	0.2789
500	3	0.2994	0.3026	0.3061	0.3097	0.3123	0.3179	0.2992	0.2609	0.2749
	6	0.2961	0.2992	0.3026	0.3062	0.3088	0.3141	0.2957	0.2573	0.2720
	9	0.2928	0.2960	0.2996	0.3033	0.3061	0.3115	0.2936	0.2556	0.2708

Table 2: RMSE of autocorrelations for FD detrending using alternative difference parameters d and lags j.

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 $\widehat{\rho}_{OLS}(j)$. Theorem 1 shows that the bias of $\widehat{\rho}_{OLS}(j)$ 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 $\widehat{\rho}_{OLS}(j)$ with $\widehat{\rho}_{v}(j)$ 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 $\widehat{\rho}_{v}(j)$ and $\widehat{\rho}_{OLS}(j)$ 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 β 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 β. For the case of IBM with only small differences in estimated β coefficients, also the differences of ACF is small. On the other extreme is Homedepot with a huge difference in β 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 $\widehat{\rho}_{OLS}(j) \xrightarrow{p} \rho_v(j)$ 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 $\widehat{\rho}_{OLS}(j)$ 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 $\widehat{\rho}_{FD}(j)$ 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 $\widehat{\rho}_{OLS}(j)$. We note that

$$\widehat{\beta}_{OLS} - \beta = \frac{\sum_{t=1}^{T}(v_t - \overline{v})(t - \overline{t})}{\sum_{t=1}^{T}\left(t - \overline{t}\right)^2} = \frac{\sum_{t=1}^{T} t v_t - \frac{1}{T}\left(\sum_{t=1}^{T} t\right)\left(\sum_{t=1}^{T} v_t\right)}{\sum_{t=1}^{T}\left(t - \overline{t}\right)^2} = O_p\left(T^{d-1.5}\right),$$

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

$$\widehat{\rho}_{OLS}(j) = \frac{\frac{1}{T}\sum_{t=j+1}^{T} (e_{t,OLS})(e_{t-j,OLS})}{\frac{1}{T}\sum_{t=1}^{T} e_{t,OLS}^2}.$$

9

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

$$
\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} e_{t,OLS}^2 &= \frac{1}{T}\sum_{t=1}^{T}(y_t - \widehat{\alpha}_{OLS} - \widehat{\beta}_{OLS}t)^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}(v_t - \overline{v})^2 - \frac{1}{T}\left(\widehat{\beta}_{OLS} - \beta\right)^2 \sum_{t=1}^{T}\left(t - \overline{t}\right)^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}(v_t - \overline{v})^2 + O_p\left(T^{2d-1}\right) \\
&= \gamma_v(0) + O_p\left(T^{2d-1}\right) \xrightarrow{p} \gamma_v(0)
\end{aligned}
$$

by Lemma 1, item 6 of Tsay and Chung (2000) and the preceding results.

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

$$
\begin{aligned}
\sum_{t=j+1}^{T}\left(e_{t,OLS}\right)\left(e_{t-j,OLS}\right) &= \sum_{t=j+1}^{T}\left(v_t - \overline{v}\right)\left(v_{t-j} - \overline{v}\right) \\
&\quad - \left(\widehat{\beta}_{OLS} - \beta\right)\sum_{t=j+1}^{T}\left(v_t - \overline{v}\right)\left(t - j - \overline{t}\right) \\
&\quad - \left(\widehat{\beta}_{OLS} - \beta\right)\sum_{t=j+1}^{T}\left(v_{t-j} - \overline{v}\right)\left(t - \overline{t}\right) \\
&\quad + \left(\widehat{\beta}_{OLS} - \beta\right)^2 \sum_{t=j+1}^{T}\left(t - \overline{t}\right)\left(t - j - \overline{t}\right) \\
&\equiv A + B + C + D.
\end{aligned}
$$

For the term A, we note that

$$
\frac{1}{T}\sum_{t=j+1}^{T}\left(v_t - \overline{v}\right)\left(v_{t-j} - \overline{v}\right) = \gamma_v(j) + O_p\left(T^{2d-1}\right) \xrightarrow{p} \gamma_v(j), \tag{2}
$$

by Theorem 3 of Hosking (1996).

For the term B, first we have

$$
\sum_{t=j+1}^{T}\left(v_t - \overline{v}\right)\left(t - j - \overline{t}\right) = \sum_{t=j+1}^{T}\left(v_t - \overline{v}\right)\left(t - \overline{t}\right) - \sum_{t=j+1}^{T} j\left(v_t - \overline{v}\right) = O_p\left(T^{d+1.5}\right), \tag{3}
$$

because $j/T \to 0$ as $T \to \infty$. This implies that $B = O_p\left(T^{d-1.5}\right)O_p\left(T^{d+1.5}\right) = O_p\left(T^{2d}\right)$.

10

For the term C, similarly, we have

$$\sum_{t=j+1}^{T} \left(v_{t-j} - \overline{v}\right)\left(t - \overline{t}\right) = \sum_{t=j+1}^{T} t v_{t-j} - \sum_{t=j+1}^{T} t\overline{v} - \sum_{t=j+1}^{T} v_{t-j}\overline{t} + \sum_{t=j+1}^{T} \overline{t}\overline{v}$$

$$= \sum_{t=j+1}^{T} t v_{t-j} + O_p\left(T^{d+1.5}\right) + O_p\left(T^{d+1.5}\right) + O_p\left(T^{d+1.5}\right).$$

Furthermore, we also note that

$$\sum_{t=j+1}^{T} t v_{t-j} = \sum_{t=1}^{T-j} t v_t + j \sum_{t=1}^{T-j} v_t = O_p\left(T^{d+1.5}\right), \tag{4}$$

and we have

$$\sum_{t=j+1}^{T} \left(v_{t-j} - \overline{v}\right)\left(t - \overline{t}\right) = O_p\left(T^{d+1.5}\right). \tag{5}$$

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

For the term D, we note that

$$\sum_{t=j+1}^{T} \left(t - \overline{t}\right)\left(t - j - \overline{t}\right) = \sum_{t=j+1}^{T} \left(t - \overline{t}\right)\left(t - \overline{t}\right) - j \sum_{t=j+1}^{T} \left(t - \overline{t}\right)^2 = O\left(T^3\right), \tag{6}$$

and we prove that $D = O\left(T^{2d-3}\right) O_p\left(T^3\right) = O_p\left(T^{2d}\right)$. Combining the asymptotic properties of A, B, C, and D, we prove that

$$\frac{1}{T}\sum_{t=j+1}^{T} \left(e_{t,OLS}\right)\left(e_{t-j,OLS}\right) = \frac{1}{T}\sum_{t=j+1}^{T} \left(v_t - \overline{v}\right)\left(v_{t-j} - \overline{v}\right) + O_p\left(T^{2d-1}\right)$$

$$= \gamma_v(j) + O_p\left(T^{2d-1}\right) \xrightarrow{p} \gamma_v(j).$$

Given the asymptotic properties of the numerator and the denominator of $\widehat{\rho}_{OLS}(j)$, we have using Slutsky's theorem that

$$\widehat{\rho}_{OLS}(j) = \frac{\gamma_v(j) + O_p\left(T^{2d-1}\right)}{\gamma_v(0) + O_p\left(T^{2d-1}\right)} \xrightarrow{p} \rho_v(j). \tag{7}$$

We now consider the asymptotic properties of $\widehat{\rho}_{FD}(j)$. Note that

$$\triangle y_t = \beta + \triangle v_t.$$

For the FD estimator, we have

$$\widehat{\beta}_{FD} - \beta = \frac{1}{T-1}\sum_{t=2}^{T} \triangle v_t = O_p\left(T^{-1}\right),$$

11

by Theorem 3 of Tsay (2000). Moreover, rewrite $\widehat{\rho}_{FD}(j)$ as

$$\widehat{\rho}_{FD}(j) = \frac{\dfrac{1}{T}\displaystyle\sum_{t=j+1}^{T}(e_{t,FD}-\overline{e}_{FD})(e_{t-j,FD}-\overline{e}_{FD})}{\dfrac{1}{T}\displaystyle\sum_{t=1}^{T}(e_{t,FD}-\overline{e}_{FD})^{2}}.$$

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

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}(e_{t,FD}-\overline{e}_{FD})^{2} &= \frac{1}{T}\sum_{t=1}^{T}(y_{t}-\widehat{\beta}_{FD}t-\overline{y}+\widehat{\beta}_{FD}\overline{t})^{2}\\
&= \frac{1}{T}\sum_{t=1}^{T}(v_{t}-\overline{v})^{2}\\
&\quad -\frac{2}{T}\left(\widehat{\beta}_{FD}-\beta\right)\sum_{t=1}^{T}(v_{t}-\overline{v})(t-\overline{t})+\frac{1}{T}\left(\widehat{\beta}_{FD}-\beta\right)^{2}\sum_{t=1}^{T}(t-\overline{t})^{2}\\
&= \frac{1}{T}\sum_{t=1}^{T}(v_{t}-\overline{v})^{2}+O_{p}(T^{d-0.5})+O_{p}(1).
\end{aligned}$$

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

$$\begin{aligned}
\sum_{t=j+1}^{T}(e_{t,FD}-\overline{e}_{FD})(e_{t-j,FD}-\overline{e}_{FD}) &= \sum_{t=j+1}^{T}(v_{t}-\overline{v})(v_{t-j}-\overline{v})\\
&\quad -\left(\widehat{\beta}_{FD}-\beta\right)\sum_{t=j+1}^{T}(v_{t}-\overline{v})(t-j-\overline{t})\\
&\quad -\left(\widehat{\beta}_{FD}-\beta\right)\sum_{t=j+1}^{T}(v_{t-j}-\overline{v})(t-\overline{t})\\
&\quad +\left(\widehat{\beta}_{FD}-\beta\right)^{2}\sum_{t=j+1}^{T}(t-\overline{t})(t-j-\overline{t})\\
&\equiv A^{*}+B^{*}+C^{*}+D^{*}.
\end{aligned}$$

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

$$\frac{1}{T}\sum_{t=j+1}^{T}(v_{t}-\overline{v})(v_{t-j}-\overline{v})=\gamma_{v}(j)+O_{p}\left(T^{2d-1}\right)\xrightarrow{p}\gamma_{v}(j).$$

For the term B^{*}, we note that

$$\sum_{t=j+1}^{T}(v_{t}-\overline{v})(t-j-\overline{t})=O_{p}\left(T^{d+1.5}\right),$$

12

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

For the term C^*, we note that

$$\sum_{t=j+1}^{T} \left(v_{t-j} - \overline{v}\right)\left(t - \overline{t}\right) = O_p\left(T^{d+1.5}\right),$$

by Equation (5), and $C^* = O_p\left(T^{-1}\right)O_p\left(T^{1.5+d}\right) = O_p\left(T^{d+0.5}\right)$.

For the term D^*, we note that

$$\sum_{t=j+1}^{T} \left(t - \overline{t}\right)\left(t - j - \overline{t}\right) = O\left(T^3\right),$$

by Equation (6). Therefore, $D^* = O_p\left(T^{-2}\right)O\left(T^3\right) = O_p\left(T\right)$. Combining the asymptotic properties of A^*, B^*, C^*, and D^*, we have

$$\frac{1}{T}\sum_{t=j+1}^{T}\left(e_{t,FD} - \overline{e}\right)\left(e_{t-j,FD} - \overline{e}\right) = \gamma_v(j) + O_p(1).$$

Finally, given the asymptotic properties of the numerator and the denominator of $\widehat{\rho}_{FD}(j)$, we have shown that

$$\widehat{\rho}_{FD}(j) = \frac{\gamma_v(j) + O_p(1)}{\gamma_v(0) + O_p(1)}. \tag{8}$$

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

References

Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5-59.

Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods, 2nd Edn. Springer, New York.

Canjels, E., Watson, M.W., 1997. Estimating deterministic trends in the presence of serially correlated errors. *Journal of Business & Economic Statistics* 79, 184-200.

Cheung, Y.-W., Lai, K.S., 1993. A fractional cointegration analysis of purchasing power parity. Journal of Business and Economic Statistics 11, 103-112.

Chipman, J.S., 1979. Efficiency of least squares estimation of linear trend when residuals are autocorrelated. Econometrica 47, 115-128.

Davydov, Y.A., 1970. The invariance principle for stationary processes. Theory of Probability and its Applications 15, 487-498.

Gallant, A.R., P.E. Rossi and G. Tauchen, 1992. Stock prices and volume, *Review of Financial Studies*, 5 , 199-242.

Granger, C.W.J., 1980. Long memory relationships and the aggregation of dynamic models. Journal of Econometrics 14, 227-238.

Granger, C.W.J., 1981. Some properties of time series data and their use in econometric model specification. Journal of Econometrics 16, 121-130.

Granger, C.W.J., 1986. Developments in the study of cointegrated economic variables. Oxford Bulletin of Economics and Statistics 48, 213-228.

Granger, C.W.J., Joyeux, R., 1980. An introduction to long-memory time series models and fractionally differencing. Journal of Time Series Analysis 1, 15-29.

Hosking, J.R.M., 1981. Fractional differencing. Biometrika 68, 165-176.

Hosking, J.R.M., 1984. Modeling persistence in hydrological time series using fractional differencing. Water Resources Research 20, 1898-1908.

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.

Krämer, W., 1982. Note on estimating linear trend when residuals are autocorrelated. Econometrica 50, 1065-1067.

Lee, C.C., Phillips, P.C.B., 1994. Efficiency gains using GLS over OLS under nonstationarity. Working paper, Yale University.

Maeshiro, A., 1976. Autoregressive transformation, trended independent variables and autocorrelated disturbance term. The Review of Economics and Statistics 58, 497-500.

McLeod, A.I., Hipel, K.W., 1978. Preservation of the rescaled adjusted range. Water Resources Research 14, 491-518.

Tauchen, G.E., Pitts, M., 1983. The price variability-volume relationship on speculative markets. Econometrica 51, 485-505.

Tsay, W.J., 2000. Estimating trending variables in the presence of fractionally integrated errors. Econometric Theory 16, 324-346.

Tsay, W.J., Chung, C.-F., 2000. The spurious regression of fractionally integrated processes. Journal of Econometrics 96, 155-182.

Yajima, Y., 1988. On estimation of a regression model with long-term stationary errors. Annals of Statistics 16, 791-807.

Yajima, Y., 1991. Asymptotic properties of the LSE in a regression model with long-memory stationary errors. Annals of Statistics 19, 158-177.

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

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 β_{OLS} and β_{FD} are the estimates of the trend coefficient using OLS and FD detrending. Standard errors are in parentheses.*
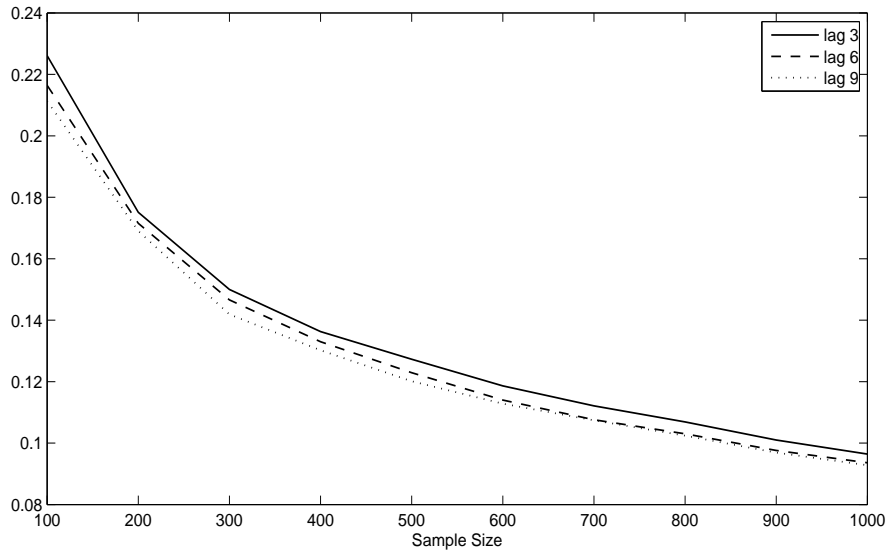
Figure 1: *Root mean squared error of ACF estimates using OLS detrending.*
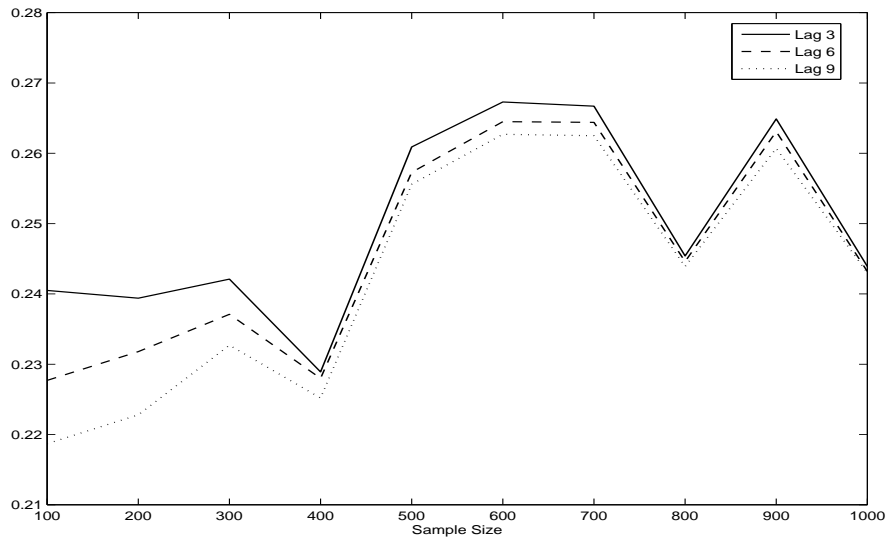

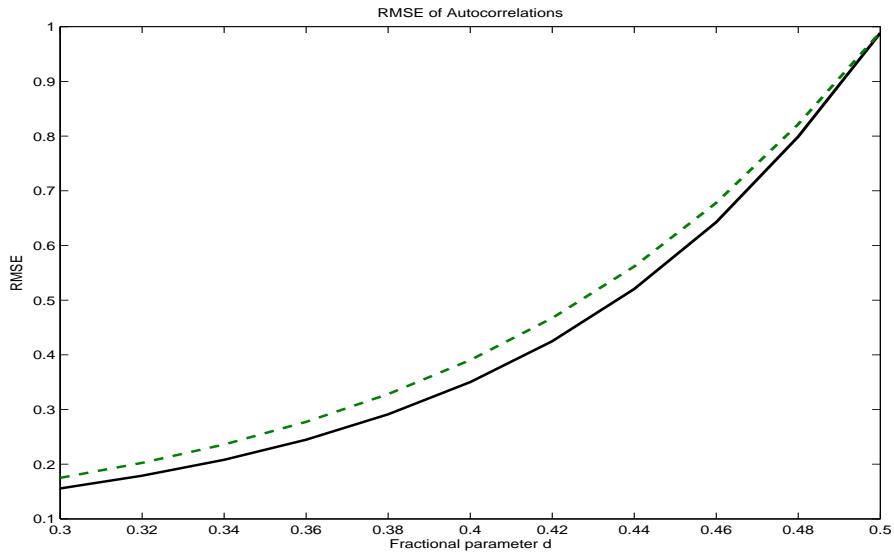Figure 2: *Root mean squared error of ACF estimates using FD detrending.*

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.*
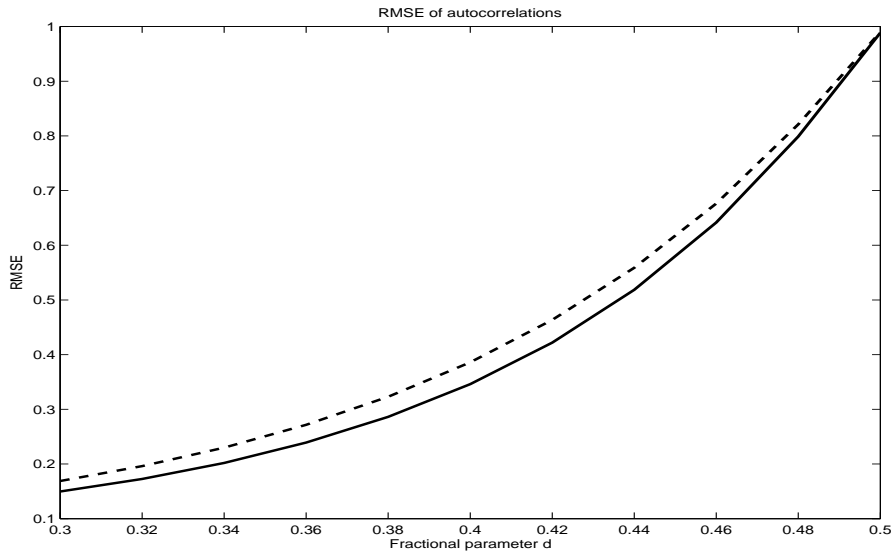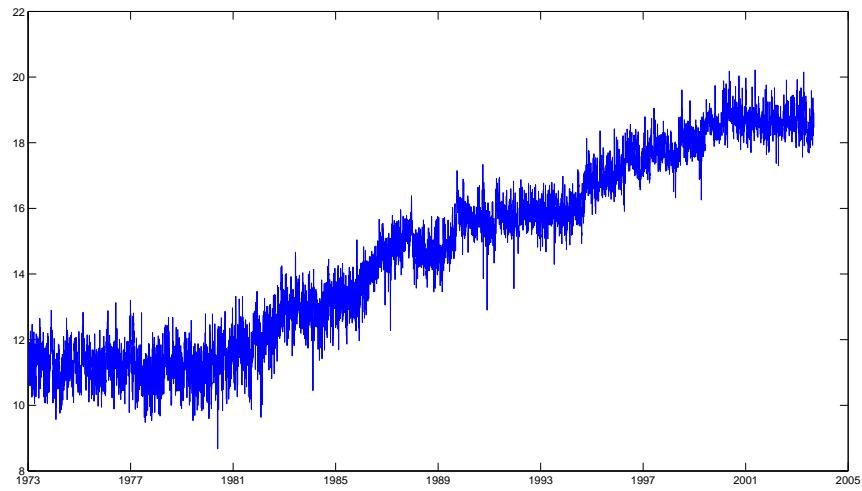

Figure 4: *RMSE of autocorrelations estimated using OLS detrending (solid line) and the true simulated values of v_t (dashed line) for a sample size of $T = 200$ and lag 9.*

Figure 5: *Time series of McDonalds daily log turnover, 1973 to 2003.*
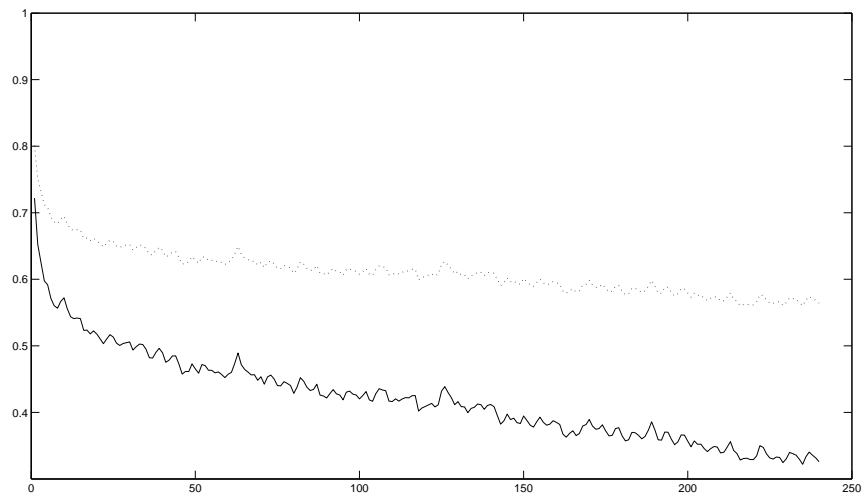
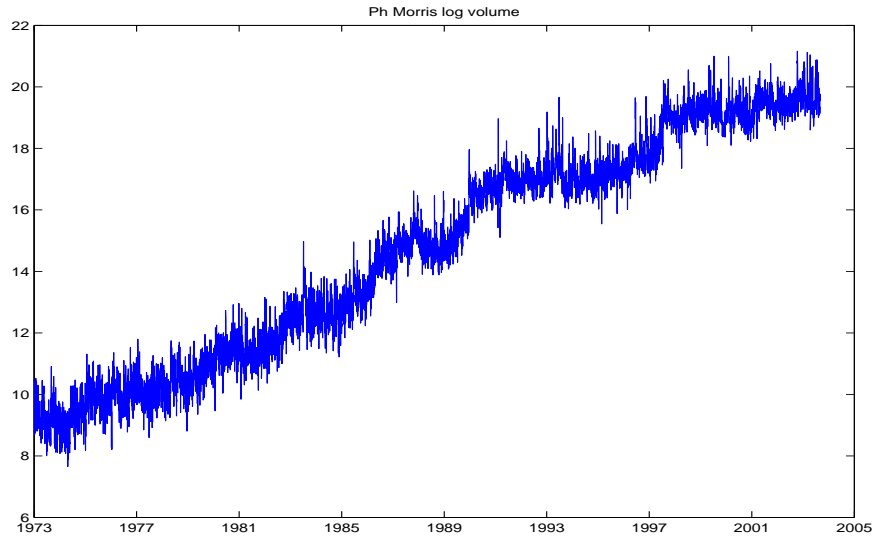Figure 6: *ACF of log turnover of McDonalds using OLS (solid) and FD (dashed) detrending.*

19

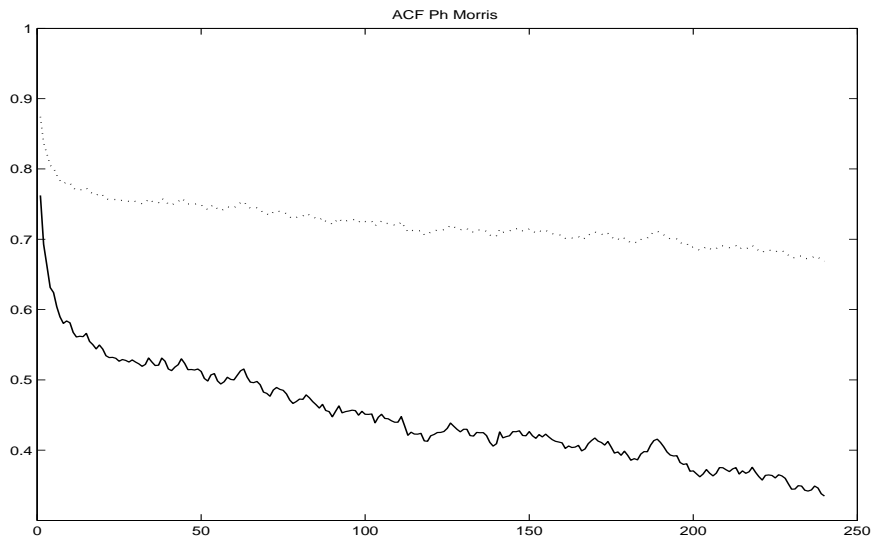Figure 7: *Time series of Philip Morris daily log turnover, 1973 to 2003.*

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

20

Recent titles

CORE Discussion Papers

2008/35. David DE LA CROIX and Clara DELAVALLADE. Democracy, rule of law, corruption incentives and growth.

2008/36. Jean J. GABSZEWICZ and Joana RESENDE. Uncertain quality, product variety and price competition. 2008/37. Gregor ZOETTL. On investment decisions in liberalized electricity markets: the impact of price caps at the spot market.

2008/38. Helmuth CREMER, Philippe DE DONDER, Dario MALDONADO and Pierre PESTIEAU. Habit formation and labor supply.

2008/39. Marie-Louise LEROUX and Grégory PONTHIERE. Optimal tax policy and expected longevity: a mean and variance approach.

2008/40. Kristian BEHRENS and Pierre M. PICARD. Transportation, freight rates, and economic geography.

2008/41. Gregor ZOETTL. Investment decisions in liberalized electricity markets: A framework of peak load pricing with strategic firms.

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.

2008/48. David DE LA CROIX. Adult longevity and economic take-off: from Malthus to Ben-Porath.

2008/49. David DE LA CROIX and Gregory PONTHIERE. On the Golden Rule of capital accumulation under endogenous longevity.

2008/50. Jean J. GABSZEWICZ and Skerdilajda ZANAJ. Successive oligopolies and decreasing returns.

2008/51. Marie-Louise LEROUX, Pierre PESTIEAU and Grégory PONTHIERE. Optimal linear taxation under endogenous longevity.

2008/52. Yuri YATSENKO, Raouf BOUCEKKINE and Natali HRITONENKO. Estimating the dynamics of R&D-based growth models.

2008/53. Roland Iwan LUTTENS and Marie-Anne VALFORT. Voting for redistribution under desert-sensitive altruism.

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.

2008/56. Santanu S. DEY. A note on the split rank of intersection cuts.

2008/57. Yu. NESTEROV. Primal-dual interior-point methods with asymmetric barriers.

2008/58. Marie-Louise LEROUX, Pierre PESTIEAU and Gregory PONTHIERE. Should we subsidize longevity?

2008/59. J. Roderick McCRORIE. 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 EYCKMANS, 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

2008/63. Jean-Sébastien TANCREZ, Philippe CHEVALIER and Pierre SEMAL. A tight bound on the throughput of queueing networks with blocking.
2008/64. Nicolas GILLIS and François GLINEUR. Nonnegative factorization and the maximum edge biclique problem.
2008/65. Geir B. ASHEIM, Claude D'ASPREMONT and Kuntal BANERJEE. Generalized time-invariant overtaking.
2008/66. Jean-François CAULIER, Ana MAULEON and Vincent VANNETELBOSCH. Contractually stable networks.
2008/67. Jean J. GABSZEWICZ, Filomena GARCIA, Joana PAIS and Joana RESENDE. On Gale and Shapley '*College admissions and stability of marriage'*.
2008/68. Axel GAUTIER and Anne YVRANDE-BILLON. Contract renewal as an incentive device. An application to the French urban public transport sector.
2008/69. Yuri YATSENKO and Natali HRITONENKO. Discrete-continuous analysis of optimal equipment replacement.
2008/70. Michel JOURNÉE, Yurii 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.
2008/73. Shin-Huei WANG and Christian M. HAFNER. Estimating autocorrelations in the presence of deterministic trends.

Books

Y. POCHET and L. WOLSEY (eds.) (2006), *Production planning by mixed integer programming.* New York, Springer-Verlag.
P. PESTIEAU (ed.) (2006), *The welfare state in the European Union: economic and social perspectives.* Oxford, Oxford University Press.
H. TULKENS (ed.) (2006), *Public goods, environmental externalities and fiscal competition.* New York, Springer-Verlag.
V. GINSBURGH and D. THROSBY (eds.) (2006), *Handbook of the economics of art and culture.* Amsterdam, Elsevier.
J. GABSZEWICZ (ed.) (2006), La différenciation des produits. Paris, La découverte.
L. BAUWENS, W. POHLMEIER and D. VEREDAS (eds.) (2008), *High frequency financial econometrics: recent developments*. Heidelberg, Physica-Verlag.
P. VAN HENTENRYCKE and L. WOLSEY (eds.) (2007), *Integration of AI and OR techniques in constraint programming for combinatorial optimization problems*. Berlin, Springer.

CORE Lecture Series

C. GOURIÉROUX and A. MONFORT (1995), Simulation Based Econometric Methods.
A. RUBINSTEIN (1996), Lectures on Modeling Bounded Rationality.
J. RENEGAR (1999), A Mathematical View of Interior-Point Methods in Convex Optimization.
B.D. BERNHEIM and M.D. WHINSTON (1999), Anticompetitive Exclusion and Foreclosure Through Vertical Agreements.
D. BIENSTOCK (2001), Potential function methods for approximately solving linear programming problems: theory and practice.
R. AMIR (2002), Supermodularity and complementarity in economics.
R. WEISMANTEL (2006), Lectures on mixed nonlinear programming.