Toward Harmless Detrending

Takeshi Otsu

Abstract

This paper investigates effects of the Christiano-Fitzgerald (CF) detrending procedure, following Chan, Hayya, and Ord (1977) and Nelson and Kang (1981). The main findings are as follows. First, inclusion of an intercept term into detrending models creates artifactual cyclical components in the spectral density function (SDF) and introduces nonlinearity in the autocorrelation function (ACF) when the true data generating process follows the random walk. Second, the Christiano-Fitzgerald (CF)-type procedure with no intercept does not create artifactual cyclical components over the frequency range corresponding to the periods of cycles observable in sample. Finally, when the CF procedure is implemented with the mean-corrected time trend, the ACF of the transformed data does not show artifactual autocorrelations, and produces a good approximation to that of the true process. These findings lead to the conclusion that the CF procedure is harmless in the linear time trend removal in that it would create less distortion in SDF and ACF of the transformed data.

1 Introduction

Many economic time-series data show trends in the sense that they move predominantly upward or downward as time goes on. Statistically, they would be well described as nonstationary: the first and/or the second moments may depend on time, and could be infinite. Since many statistical tools are built on the assumption of stationarity, researchers, engaged in applied work, attempt to stationarize data by some detrending methods. In empirical economics, the regression-based linear time trend removal and the first differences are the most popular detrending methods. The regression residuals or the differenced data are supposed to be trend-free and even stationary. These methods, however, might be used inappropriately because it is not known a priori what statistical model the true trend components follow. Chan, Hayya, and Ord (1977) examined the spectral density functions (SDF) and the autocorrelation functions (ACF) of the simple linear regression residuals when the true data−generating process is a random walk, the summation of a purely random series, and those of the first differenced data when the true process is a linear trend plus a random series. They found that a wrong choice of model creates artifacts in the SDF and the ACF: reweighting in SDF and spurious values of ACF. The wrongly chosen linear regression model produces exaggeration at the low frequency portion of the SDF and large positive autocorrelations in the first few lags, while the inappropriately differenced series show exaggeration at the high frequency and a negative autocorrelation (−0.5) at the first lag. Nelson and Kang (1981) also found spurious dynamic properties of residuals from regression of a random walk on time: the nonlinearity of autocorrelations and a primary peak in the SDF at a period equal to 0.80 of the sample size. Further, Nelson and Kang (1984) studied effects on the statistical inference in the regression analysis, and found that the spurious dynamics of residuals from regression with a time trend as an explanatory variable produced a statistically significant relationship between the explained and the explanatory variables where none exists. They recommended estimation in differenced form, because over-differencing still preserves consistency of least−squares estimators, though inefficient.

In this paper, we seek a trend-removal method that would create less distortion in the SDF and the ACF of the transformed data. This implies that the transformed data would be nonstationary as long as the original is. That is, our aim of a trend removal is not to stationarize data but to simply remove a linear trend. The motivation comes from a general property of the digital filtering methods. With the theoretical background of the signal extraction based on the Wiener-Kolmogorov theory (Whittle, 1983, Chapter 3 and 6), the filtering methods are widely used in empirical analyses of macroeconomics: estimation of potential output, business cycles, and seasonal adjustments, as briefly discussed in Otsu (2010). The filtering methods, in general, assume stationarity just as other statistical tools. Bell (1984), however, showed that the Wiener-Kolmogorov theory is applicable to nonstationary series as long as the series are differenced-stationary and the initial values are generated independently of the differenced data (see Bell, 1984, p. 652). Thus, as long as the data are difference-stationary, the filtering analyses would be useful. The remaining problems are how to remove a linear trend without distorting the original statistical property and to make valid the circularity assumption in filtering discussed below.

A predominant time trend typical in macroeconomic time series, however, may prevent adequate use of filtering in practice. We can implement filtering either in the time domain or in the frequency domain. But, the time-domain filtering would introduce phase shifts in transformed data unless we can exclude endpoints of data from economic analyses as many as the filter lengths. Therefore, the frequency-domain filtering would be preferred to preserve the original phase, as argued in Iacobucci and Noullez (2005) and Otsu (2009). To implement filtering in the frequency domain, it is typical to use the discrete Fourier transform, which requires the circularity of data: the first observation can be considered as the next of the last observation. When a time trend is observed, the starting value would be very different from the ending value; then, the discrepancy between both ends of the data causes serious distortion in transformation. Then, it is required that the trend removal should make the circularity assumption tenable. If the circularity is guaranteed, the differencing would take care of the remaining nonstationariy in data. So long as the trend removal does not create distortion in the SDF and the ACF, the subsequent analyses would be appropriately conducted. The detrending procedure used in Christiano and Fitzgerald (2003) would be a promising candidate for a harmless trend removal, because it makes equal both endpoints of the transformed data.

To inspect detrending effects, we look at autocorrelation functions (ACF) and spectral density functions (SDF). We derive the expected sample ACF and the corresponding SDF in a manner of Chan, Hayya, and Ord (1977) and Nelson and Kang (1981). Then, we compare the ACF's and the SDF's values of the original series against those of the trend-removed series via the ordinary least squares (OLS) and the Christiano-Fitzgerald (CF) procedure used in Christiano and Fitzgerald (2003). The main findings are as follows. First, inclusion of an intercept term into a detrending model creates artifactual cyclical components in the spectral density function (SDF) and introduces nonlinearity in the autocorrelation function (ACF) when the true data generating process follows the random walk. Second, the Christiano-Fitzgerald (CF)-type procedure with no intercept does not create artifactual cyclical components over the frequency range corresponding to the periods of cycles observable in sample. Finally, when the CF procedure is implemented with the mean-corrected time trend, the ACF's values of the transformed data are not seriously distorted, and produces a good approximation to those of the true process. These findings lead to the conclusion that the CF procedure is harmless in the linear time trend removal.

The remaining part of this paper goes as follows. In section 2, we discuss the Christiano Fitzgerald (CF) drift-adjusting procedure. In section 3, we derive the expected sample autocorrelation functions and the spectral density functions of the series that are supposed to be detrended by the OLS, the CF procedure and their variants, when the true process is a random walk. Section 4 is allocated for the case that the true process follows the trend stationary. We conduct graphical and quantitative examinations of the derived ACFs and the corresponding SDFs to inspect possible distortion by trend removals. Final remarks are given in section 5.

2 Christiano-Fitzgerald Detrending

Christiano and Fitzgerald (2003, p. 439) proposed the following drift adjusting procedure when they use their bandpass filters. Let the raw data X_t , $t = 1, \cdots, N$. Then, we compute the drift-adjusted series, w_t, as follows:

$$w_t = X_t - (t - 1)\hat{\beta}_1 \tag{1}$$

where

$$\hat{\beta}_1 = \frac{X_N - X_1}{N - 1} \tag{2}$$

Note that $w_1 = w_N (= X_1)$. Thus, the drift-adjusting procedure in eq.(1) would make the data suitable for filtering in the frequency domain. The presumed data-generating process of X_t is

$$X_t = \beta_1 + X_{t-1} + \varepsilon_t \tag{3}$$

where β_1 indicates a constant drift term, and ε_t follows a zero-mean covariance-stationary process. That is,

$$E(\varepsilon_t) = 0 \tag{4}$$

$$E(\varepsilon_t \varepsilon_{t-s}) = \begin{cases} \sigma^2 & \text{if } s = 0 \\ 0 & \text{if } s \neq 0 \end{cases} \tag{5}$$

Further, the spectrum of ε_t is, by normalizing,

$$SDF_{\varepsilon}(f) = 1, \qquad 0 \leq f \leq \frac{1}{2} \tag{6}$$

where f is the frequency. Then, as pointed out by Christiano and Fitzgerald (2003), X_t can be expressed as

$$X_t = (t + q)\beta_1 + w_t \tag{7}$$

where q is a fixed integer, and

$$w_t = w_{t-1} + \varepsilon_t \quad \text{for all } t \tag{8}$$

It can be seen that a linear time trend is implicitly embedded in eq.(3). Christiano and Fitzgerald (2003) set q to -1, to obtain eq.(1). The choice of q simply changes the level of the transformed data, and it can be any integer. Since the transformation from eq.(3) to eq.(8) removes the drift, β_1, this procedure adjusts a drift by removing a linear line in the unit root case. This suggests the trend removal by eq.(1) preserve statistical properties of the original series in the transformed series. Further, the estimator in eq.(2) can be viewed as a least squares estimator in a trivial sense. We rewrite eq.(3) as

$$X_t - X_{t-1} = \beta_1 + \varepsilon_t \tag{9}$$

Then, we can estimate β_1 by regressing $X_t - X_{t-1}$ on to a constant term, which results in eq.(2). Just as the ordinary least squares estimator is unbiased, so is this estimator. If we use the estimate given by eq.(2) in eq.(7) and subtract the first term from X_t, the remaining components should follow the process in eq.(8). Therefore, we successfully remove the drifting term in eq.(3) or the time-trend component and the constant term in eq.(7).

Now, suppose the true data−generating process is as follows:

$$X_t = \beta_0 + \beta_1 t + \varepsilon_t \tag{10}$$

Further, we manipulate this to obtain

$$X_t - X_{t-1} = \beta_1 + \varepsilon_t - \varepsilon_{t-1} \tag{11}$$

Therefore, the least squares method gives rise to the unbiased estimator of β_1 as before, although not efficient. Then, it would be appropriate to use the estimate of eq.(2) as a slope coefficient on a time trend in eq.(10). The statistical property of the computed residuals would be completely determined by ε_t. The trend components in eq.(10) might be successfully removed.

To sum up, the trend removal with a slope estimated by eq.(2) would be harmless in the sense that the detrending does not change statistical properties of the original series, whether the data generating process follows a unit root process or a trend-stationary process. In other words, we can get rid of nonstationarity purely attributable to a linear time trend. If we succeed to remove a trend without distorting statistical properties, we conjecture that the nonstationarity left behind can be dealt with differencing. This would make applied researchers' life easier, since differencing would make useful the conventional statistical tools assuming stationarity and over differencing would be less costly, according to the conventional wisdom.

This presents a striking contrast with the findings of Chan, Hayya, and Ord (1977) and Nelson and Kang (1981) about the OLS-type trend removal when the true data generating process follows a unit root process. In the following sections, we inspect the expected sample ACF and the corresponding SDF of the series detrended with $\hat{\beta}_1$ in eq.(2). We focus on a finite sample case, because sample is always finite in practice.

3 Trend Removal with the Random Walk Process

Suppose that the true data generating process follows a random walk with a drift as in eq.(3). Following Chan, Hayya, and Ord (1977), we consider the sample size of $N = 2n + 1$, and the time index, t, goes from –n through zero to n. Then, we have zero sample mean of the time trend:

$$\bar{t} = \frac{1}{N}\sum_{t=-n}^{n} t = 0 \tag{12}$$

where $N = 2n + 1$ We further assume $X_{-n-1} = 0$. Now, suppose we use a linear time-trend model to remove a trend in the original series. That is, we presume that the specification in eq.(10) is appropriate. Chan, Hayya, and Ord (1977) and Nelson and Kang (1981) examined the case that the OLS estimates of β_0 and β_1 were used to remove the trend. The OLS estimators are given by

$$\hat{\beta}_0 = \overline{X}, \qquad \text{(the sample mean of } X\text{)} \tag{13}$$

$$\hat{\beta}_{OLS} = \frac{\sum_{k=-n}^{n} kX_k}{\sum_{k=-n}^{n} k^2} \tag{14}$$

Then, the OLS residuals are supposed trend-free:

$$e_{\mathrm{t}} = X_{\mathrm{t}} - \hat{\beta}_0 - \hat{\beta}_{\mathrm{OLS}} t \tag{15}$$

To inspect the statistical properties, we carry out iterative substitutions in eq.(3) to obtain

$$X_{\mathrm{t}} = \sum_{\mathrm{k}=-\mathrm{n}}^{\mathrm{t}} \varepsilon_{\mathrm{t}} + \beta_1 (t + n + 1) \tag{16}$$

Using eq.(13), eq.(14), and eq.(16), it is easy to rewrite eq.(15) as

$$\begin{aligned} e_t &= \sum_{k=-n}^{t} \varepsilon_k - \frac{1}{N} \sum_{m=-n}^{n} \sum_{k=-n}^{m} \varepsilon_k \\ &\quad - \frac{t}{\sum_{k=-n}^{n} k^2} \left\{ \sum_{m=-n}^{n} \sum_{k=-n}^{m} \varepsilon_k \right\} \end{aligned} \tag{17}$$

Then, the theoretical autocovariance functions with a lag length (s) is

$$\begin{aligned} cov(e_{\mathrm{t}}, e_{\mathrm{t-s}}) &= \frac{\sigma_2}{30n(n+1)(2n+1)} \{10n^2(n+1)^2 \\ &\quad - 15n(n+1)(2n+1)s + 15n(n+1)s^2 \\ &\quad + (24n^2 + 24n + 12)st - 15s^3 t \\ &\quad - (24n^2 + 24n + 12)t^2 + 45s^2 t^2 \\ &\quad - 60\, st^3 + 30t^4\} \end{aligned} \tag{18}$$

In practice, we estimate autocovariances by averaging over the time interval of the sample: summing from $-n+s$ to n and divided by N. Then, following Chan, Hayya, and Ord (1977), we have the expected sample autocovariance function as in Nelson and Kang (1981, p. 742). Then, we obtain the expected sample autocorrelation function, dividing eq.(18) by the variance, that is, the covariance when $s = 0$:

$$\begin{aligned}\bar{\rho}_{\mathrm{e}}(s,n) = 1 + \{ &-(152n^4 + 304n^3 + 150n^2 - 2n - 6)s \\ &+ (180n^3 + 270n^2 + 90n)s^2 \\ &-(68n^2 + 68n + 9)s^3 \\ &+ 3s^5\}\frac{1}{32n^5 + 80n^4 + 40n^3 - 20n^2 - 12n}\end{aligned} \tag{19}$$

To compute the corresponding spectral density function (SDF), we use the following formula for some autocorrelation function $\bar{\rho}(s,n)$:

$$SDF(f) = 1 + 2\sum_{s=1}^{2n}\bar{\rho}(s,n)cos(2\pi sf), \qquad 0 \le f \le \frac{1}{2} \tag{20}$$

We also consider the case of zero intercept restriction. The autocorrelation function is similarly derived as follows. That is, the detrended series are computed as

$$u_{\mathrm{t}} = X_{\mathrm{t}} - \hat{\beta}_{\mathrm{OLS}}t \tag{21}$$

Then, the corresponding autocorrelation function is

$$\begin{aligned}\bar{\rho}_{\mathrm{u}}(s,n) = 1 + \{ &-(64n^4 + 148n^3 + 100n^2 + 16n - 2)s \\ &+ 20n(n+1)(2n+1)s^2 \\ &-(16n^2 + 16n + 3)s^3 \\ &+ s^5\}\frac{1}{4n(n+1)(2n+1)(8n^2 + 13n + 4)}\end{aligned} \tag{22}$$

Now, we use the slope estimator proposed by Christiano and Fitzgerald (2003):

$$\hat{\beta}_{\mathrm{CF}} = \frac{X_{\mathrm{n}} - X_{-\mathrm{n}}}{2n} \tag{23}$$

and $\hat{\beta}_0$ is given by eq.(13). Let z_t the detrended series from eq.(15). Note that, in contrast to the OLS-based residuals, the transformed series have the same values at both endpoints. That is,

$$z_{-n} = z_n = \frac{X_n + X_{-n}}{2} - \overline{X} \tag{24}$$

This property makes the Fourier transform suitable in filtering. The autocorrelation function is given by

$$\begin{aligned}\overline{\rho}_z(s,n) = 1 + \{ &-(28n^3 + 22n^2 - 1)s \\ &+ 12n(2n+1)s^2 \\ &-(6n+1)s^3\}\frac{1}{2n(n+1)(2n+1)(2n-1)}\end{aligned} \tag{25}$$

In comparison with eq.(19) and eq.(22), both the magnitude of coefficients and the orders of n and s are smaller, implying a lesser degree of non-linearity of the autocorrelation. We also consider detrending without an intercept, using eq.(23):

$$w(q)_t = X_t - \hat{\beta}_{CF}(t+q) \tag{26}$$

where q could be any integer as argued in section 2. When we set q to n, this is corresponding to the drift-adjusting procedure in Christiano and Fitzgerald (2003) that showed the endpoints of the transformed data are the first observation of the original series. That is,

$$w(n)_{-n} = w(n)_n = X_{-n} \tag{27}$$

The autocorrelation for this case is

$$\begin{aligned}\overline{\rho}_{w(n)}(s,n) = 1 + \{ &-(12n^2 + 12n - 1)s \\ &+ 6ns^2 - s^3\}\frac{1}{2n(2n+1)(2n+5)}\end{aligned} \tag{28}$$

The coefficients on s and its powers, in absolute values, are much smaller than those of eq.(25). If we set q to zero in eq.(26), we have

$$w(0)_t = X_t - \hat{\beta}_{CF} t \tag{29}$$

Then, the transformed series have the same values at both endpoints:

$$w(0)_{-n} = w(0)_n = \frac{X_n + X_{-n}}{2} \tag{30}$$

And we have the autocorrelation function as follows.

$$\begin{aligned}\bar{\rho}_{w(0)}(s, n) = 1 + \{ &-(18n^2 + 12n - 1)s \\ &+ 6ns^2 - s^3\} \frac{1}{10n(n+1)(2n+1)}\end{aligned} \tag{31}$$

Then, the magnitude of the coefficient on s would be slightly larger than that in eq.(28). We compare the ACF and the corresponding SDF discussed above with those of the random walk, the summation of ε_t specified in eq.(4) and eq.(5), because the transformed series should follow the random walk process if the trend removal is successfully implemented. Although the spectral density function is not well-defined for nonstationary data, it is possible to compute by eq.(20) in practice for a given n and s. Then, the autocorrelation function for the random walk can be written as

$$\bar{\rho}_{RW}(s, n) = 1 - \frac{(4n+3)s}{2(n+1)(2n+1)} + \frac{s^2}{2(n+1)(2n+1)} \tag{32}$$

This is much simpler than any other autocorrelation function discussed here. We compare the autocorrelation in eq.(32) and the corresponding SDF with those of the various detrending procedure discussed above.

Figure 1 shows the autocorrelations and the spectral density over the frequency range from 0.0 to 0.1 when the sample size (N) is 101 or equivalently n is 50. The upper panels include replications of the autocorrelation functions and the spectral density functions shown in Nelson and Kang (1984, p. 743 and p. 745). As they pointed out, we observe that the OLS-based trend removal creates a peak in the SDF at the frequency of 0.0125 or 80 periods per cycle. The autocorrelation function presents nonlinearity with negative values over the middle range. Thus, generally speaking, we should be very cautious in use of the linear trend in empirical analyses. These phenomena, however, are not peculiar to the OLS-based procedure. When we use the CF-type slope estimator, we also observe similar distortion of the SDF and the ACF. The lower panels show the SDF with a peak at the frequency of 0.01, 100 periods per cycle and the nonlinear ACF similar to that of the OLS transformation. As exemplified by Nelson and Kang (1984), such an artifactual nonlinearity of the autocorrelation functions might affect statistical inferences and mislead researchers.

When we drop the intercept term, we do not observe such a strong nonlinearity. In Figure 2, the autocorrelation from each method monotonically decreases as the lag, s, is getting larger, and the pseudo SDF of the finite sample looks similar each other. Particularly, when the mean-corrected time trend is used, that is, $\bar{t} = 0$, both the OLS-based and the CF-based trend removal give very similar ACF and SDF: in both cases, the autocorrelations take larger values over the middle lag lengths than those of the pure random walk process. Although the non-zero-mean time trend case, that is, the CF detrending with $q = n$, gives rise to smaller autocorrelations and smaller density near zero frequencies, the general shapes still look similar. Therefore, a large part of nonlinearity comes from the effect of the mean correction through the intercept term that suppresses the spectrum at zero frequency.

We also look at some quantity measures to inspect the degree of distortion. Table 1 presents approximate theoretical values of the SDF at zero and 0.5 frequencies. If we draw the pseudo SDF with random-walk data, the density is approximately $4n/3$ at zero frequency and $1/(2n)$ at the highest frequency. The models with an intercept give a larger density at 0.5 frequency than the random-walk data: five times larger for the OLS and four times for the CF, respectively. Without an intercept term, the OLS produces a larger density by a factor of 1.25 at both 0.0 and 0.5 frequencies, compared with the pseudo spectral density of the random walk process. Similarly, the CF gives a larger density by a factor of 1.20 at both frequencies when q is set to zero. When q is set to n, however, it gives a smaller value of the density by a factor of 0.75 at zero frequency and a larger value by a factor of 1.5 at 0.5 frequency. Thus, the CF-type detrending with the mean-corrected time trend seems less distortionary.

The large deviations in the SDF exist near zero frequency, as observed in Figure 1 and 2. Then, we have a closer look over the frequencies up to 0.0225 in Table 3. When there is no intercept term, the OLS and the CF with the parameter q set to zero produce very similar values, which match the pseudo density values of the random walk process more closely than those of the CF with q equal to n. The density of the CF $(q = n)$ gives overvalues over the range of 0.0075 to 0.0225 frequencies. Here, we note that observable cycles would be less than equal to 100 because the sample size is set to 101. This implies that the cycles longer than 100 periods would not be for analyses in practice. In other words, deviations at frequencies lower than 0.01 will not be harmful to empirical analyses. Then, the CF $(q = 0)$ would be a good detrending method because it produces reasonably appropriate values of spectral density over the frequency range greater than equal to 0.01.

To see how the autocorrelations of each transformed series deviate from those of the ideally detrended "*true*" series, we use the root mean-squared error (RMSE) as an overall measure:

$$RMSE\,(\rho_{true}(s,n),\overline{\rho}(s,n)) = \sqrt{\frac{\sum_{s=1}^{2n}(\rho_{true}(s,n)-\overline{\rho}(s,n))^2}{2n}} \qquad (33)$$

where $\rho_{\mathrm{true}}(s,n)$ is given in eq.(32) for the random walk case and takes zero for the white noise case for $s \geq 1$. $\overline{\rho}(s,n)$ takes one of the autocorrelation functions in eq.(19), eq.(22), eq.(25), eq.(28) or eq.(31). The values for each model are presented in Table 2. The smallest deviation is observed in the case of the CF with $q = 0$. Further, Table 4 shows autocorrelations at selected lags. As in the SDF, the OLS and the CF ($q = 0$) with no intercept give similar values. Over the lag lengths from 1 to 50, their autocorrelations are close to those of the random walk, while those of the CF with q set to n are closer over the lags greater than 60. Although the choice of q gives a slightly different result, it would be fair to say that the CF procedure, used with the mean-corrected time trend, performs relatively well.

4 Trend Removal with the Trend-Stationary Process

In this section, we assume that the true data-generating process follows a simple trend-stationary (TS) process in eq.(10). Then, the autocorrelations for $s > 0$ are zero, and the value of SDF defined in eq.(20) is one over all the frequency range. If we happen to use the true model to remove a trend based on the OLS estimates in eq.(13) and eq.(14), we should be able to successfully detrend the series. The detrended series are obtained through eq.(15). The autocorrelation is derived as in the previous section. Obviously, the autocorrelation, $\bar{\rho}_{\mathrm{e}}^{\mathrm{TS}}(s,n)$, takes one when s is zero. For $s>0$, the autocorrelation function is

$$\begin{aligned}\bar{\rho}_{\mathrm{e}}^{\mathrm{TS}}(s,n) = \{ &-4n(n+1)(2n+1)\\ &+(8n^2+8n+1)s\\ &-s^3\}\frac{1}{2n(n+1)(2n-1)(2n+1)}\end{aligned} \tag{34}$$

It is easily seen that the autocorrelation is asymptotically zero as n goes to infinity. If we drop the constant term and use eq.(21) with $\hat{\beta}_{\mathrm{CF}}$ instead of $\hat{\beta}_{\mathrm{OLS}}$ to remove the trend, the autocorrelation function ($s>0$) is

$$\begin{aligned}\bar{\rho}_{\mathrm{u}}^{\mathrm{TS}}(s,n) = \{ &-2n(n+1)(2n+1)\\ &+(6n^2+6n+1)s\\ &-s^3\}\frac{1}{4n^2(n+1)(2n+1)}\end{aligned} \tag{35}$$

Similarly, when the true process is described as trend-stationary and the $\hat{\beta}_{\mathrm{CF}}$ in eq.(23) is used in the trend removal, we have the following autocorrelation functions for $s>0$, instead of eq.(25), eq.(28), and eq.(31), respectively:

$$\begin{aligned}\bar{\rho}_{\mathrm{z}}^{\mathrm{TS}}(s,n) = \{&2n(2n+1)(2n^2-9n+1)\\ &-(12n^3-18n^2-4n+1)s\\ &+(2n+1)s^3\}\frac{1}{2n(2n+1)(14n^2-9n+1)}\end{aligned} \tag{36}$$

$$\begin{aligned}\bar{\rho}_{w(\mathrm{n})}^{\mathrm{TS}}(s,n) = \{&2n(2n+1)(4n+1)\\ &+(6n^2+6n+1)s\\ &+s^3\}\frac{1}{2n(2n+1)(10n+1)}\end{aligned} \tag{37}$$

$$\bar{\rho}^{TS}_{w(0)}(s,n) = \{2n(n+1)(2n+1) - (12n^2+6n+1)s + s^3\}\frac{1}{2n(2n+1)(7n+1)} \tag{38}$$

In Figure 3, we use the models with an intercept to remove a time trend. Since the successfully detrended series should follow the white noise process, we compare the ACF and the SDF of the detrended series with those of the white noise. The SDFs of the detrended series have zero value at zero frequency due to the mean-correction effect of the intercept term. The OLS-based transformation produces almost no artifactual cyclical component and flattened spectra over the frequencies above 0.125, 80 periods per cycle, which is consistent with those of the white noise. In contrast, the CF-based transformation creates modest cyclical components up to 0.03 frequencies, 33 periods per cycle, and the first peak at 133 periods per cycle. Further, the ACF of the OLS-based procedure matches that of the white noise, while the CF-based ACF shows overestimation up to 34 lags and underestimation afterwards. In sum, the OLS-based does better than the CF-based, as expected.

When we drop the intercept, the OLS-based procedure produces the SDF close to that of the white noise, as shown in the upper panels of Figure 4. In the middle panels, we find that the CF-based with the mean-corrected time trend ($\bar{t} = 0$) does well as much as the CF-based with an intercept. It also has the first peak in the SDF at 133 periods per cycle and negative autocorrelations at lags greater than 37. The bottom panels show the case that the time trend has none zero mean ($q = n$). The CF-based produces significant pseudo cyclical components over the frequencies lower than 0.01, 100 periods per cycle, and large artifactual positive autocorrelations. Note, however, we are supposed to observe cyclical components only up to the 100-period cycle because the number of sample is 101. Then, the distortions over the frequencies of these unobservable cycles would not be problems in practice.

Let us look at the quantity measures same as in the previous section. In Table 1, we find that the OLS without an intercept gives the same SDF as the white noise process at both zero and the highest (0.5) frequencies, while the CF procedure gives a smaller value by a factor of 0.86 when q is set to zero. Interestingly, it generates $3n/5$ at zero frequency and 1/5 at 0.5 frequency when q is equal to n. This SDF looks like that of a nonstationary series. Thus, it is not recommendable to use the time trend without the mean correction, though the SDF values at frequencies higher than 0.01 are comparable among the CF-type procedures in Table 5.

In terms of the overall deviation measure (RMSE) in Table 2, the OLS without an intercept does better than with an intercept. The performance of the CF procedure does not depend on the existence of the intercept, but on whether or not the time trend is mean-corrected: the deviation is substantially reduced with the mean-corrected trend. This point is also confirmed in Table 6. Without the mean-corrected trend, the autocorrelations comparatively take large positive values even at 50 lags, which can be misleading. As already seen in the previous section, the choice of the parameter q matters. In sum, if we need to make the circularity assumption tenable, the CF detrending method with the mean-corrected time trend would be the best choice.

5　Concluding Remarks

In this paper, we investigated the Christiano-Fitzgerald (CF) detrending procedure, which is a promising candidate for a harmless trend-removal method in the sense that it would not create serious distortion in the spectral density functions (SDF) and the autocorrelation functions (ACF) of the transformed data. We derived the expected sample ACF and the corresponding SDF in a manner of Chan, Hayya, and Ord (1977) and Nelson and Kang (1981). Then, we examined detrending methods based on the ordinary least squares (OLS) and the Christiano-Fitzgerald (CF) procedure used in Christiano and Fitzgerald (2003). The main findings are as follows. First, inclusion of an intercept term into detrending models creates artifactual cyclical components in the spectral density function (SDF) and introduces nonlinearity in the autocorrelation function (ACF) when the true data generating process follows the random walk. Second, the Christiano-Fitzgerald (CF)-type procedure with no intercept does not create artifactual cyclical components over the frequency range corresponding to the periods of cycles observable in sample. Finally, when the CF procedure is implemented with the mean-corrected time trend, the ACF of the transformed data does not make seriously distorted autocorrelations, and produces a good approximation to that of the true process. This is very attractive because the values of the autocorrelations would be very important to analyses or interpretation of the data, particularly in economics. These findings lead to the conclusion that the CF procedure is harmless in the linear time trend removal.

We point out the usefulness of the harmless detrending to close our argument. We conjecture that if the pure time trend is successfully removed without distortion, the remaining nonstationarity can be dealt with differencing. Further, since the Christiano-Fitzgerald (CF)-type procedure makes the circularity assumption of data tenable, the remaining nonstationarity can be handled by digital filtering via the Fourier Transform. Then, the signal extraction theory would become suitable for empirical analyses in economics.

Acknowledgements: This work contributes to part of the project under the title of「新興国の台頭による日本の産業・企業への影響ならびに再構築に向けた諸課題」(成城大学特別研究助成 2010－2011 年度). The author gratefully acknowledges the financial support from JSPS (Japan Society for the Promotion of Science) KAKENHI (17530223).

References

Bell, William. R., 1984, "Signal extraction for nonstationary time series," The Annals of Statistics 12 (2), 646-664.

Chan, Hung C, Jack C. Hayya, and J. Keith Ord, 1977, "A note on trend removal methods: The case of polynomial regression versus variate differencing," Econometrica 45, 737-744.

Christiano, Lawrence J. and Terry J. Fitzgerald, 2003, "The band pass filter," International Economic Review 44 (2), 435-465.

Iacobucci, Alessandra and Alain Noullez, 2005, "A frequency selective filter for short-length time series," Computational Economics 25, 75-102.

Nelson, Charles R. and Heejoon Kang, 1981, "Spurious periodicity in inappropriately detrended time series," Econometrica 49, 741-751.

Nelson, Charles R. and Heejoon Kang, 1984, "Pitfalls in the use of time as an explanatory variable in regression," Journal of Business and Economic Statistics 2 (1), 73-82.

Otsu, Takeshi, 2009, "Seasonal cycle and filtering," Seijo University Economic Papers 185, 175-221, The Economic Institute of Seijo University (Japan).

Otsu, Takeshi, 2010, "Separating trends and cycles," Seijo University Economic Papers 188, 95-158, The Economic Institute of Seijo University (Japan).

Whittle, Peter, 1983, Prediction and regulation by linear least-square methods, 2nd ed. (revised), Basil Blackwell, Oxford.

Table 1 Approximate Theoretical Spectral Density

Frequency	True	True Process: Random Walk (R.W.)				
		With Intercept		No Intercept		
	R.W.(*)	OLS	CF	OLS	CF (q=0)	CF (q=n)
0.0	$\frac{4n}{3}$	0	0	$\frac{5n}{3}$	$\frac{8n}{5}$	n
0.5	$\frac{1}{2n}$	$\frac{5}{2n}$	$\frac{2}{n}$	$\frac{5}{8n}$	$\frac{3}{5n}$	$\frac{3}{4n}$

Frequency	True	True Process: Trend Stationary (T.S.)				
		With Intercept		No Intercept		
	W.N.(**)	OLS	CF	OLS	CF (q=0)	CF (q=n)
0.0	1	0	0	1	$\frac{6}{7}$	$\frac{3n}{5}$
0.5	1	1	$\frac{6}{7}$	1	$\frac{6}{7}$	$\frac{1}{5}$

Note: Approximated by the leading factors.
* Pseudo values by the conventional SDF formula, given the number of samples.
** Abbreviation of the White Noise.

Table 2 Deviation of Autocorrelations

True Process	With Intercept		No Intercept		
	OLS	CF	OLS	CF (q=0)	CF (q=n)
Random Walk	0.4606	0.4249	0.0969	0.0751	0.0872
Trend Stationary	0.0084	0.0552	0.0041	0.0596	0.1932

Note: Measured by root mean−squared errors in eq. (33)

Table 3 Selected Spectral Density: Random Walk Case

Freq-uency	True	With Intercept		No Intercept		
	R.W.(*)	OLS	CF	OLS	CF (q=0)	CF (q=n)
0.0000	67.6667	0.0000	0.0000	84.3786	81.2000	52.9143
0.0025	59.5705	0.2944	3.9621	69.6468	67.7533	48.0913
0.0050	40.5237	3.7441	14.3200	37.7918	38.5825	36.3853
0.0075	21.4523	12.8243	25.7599	12.1868	14.7467	23.6725
0.0100	9.8385	23.0565	30.7172	3.6772	5.9091	14.3364
0.0125	5.5957	26.5370	25.8086	5.1107	5.7751	9.0769
0.0150	4.5054	21.4958	15.9575	5.7214	5.2299	6.3413
0.0175	3.5515	13.9982	9.1303	3.3427	2.7898	4.6901
0.0200	2.4622	9.9018	7.6942	1.6035	1.4833	3.5881
0.0225	1.8405	8.9638	7.7721	1.8128	1.8486	2.8718

* Pseudo values by the conventional SDF formula, given the number of samples.

Table 4 Selected Autocorrelations: Random Walk Case

	True	With Intercept		No Intercept		
Lag	R.W.	OLS	CF	OLS	CF (q=0)	CF (q=n)
1	0.9804	0.9081	0.9315	0.9806	0.9824	0.9714
10	0.8127	0.2633	0.4158	0.8255	0.8342	0.7388
20	0.6447	−0.1303	0.0339	0.6880	0.6894	0.5285
30	0.4962	−0.2755	−0.1812	0.5768	0.5632	0.3635
40	0.3671	−0.2602	−0.2649	0.4822	0.4533	0.2381
50	0.2574	−0.1622	−0.2526	0.3962	0.3574	0.1467
60	0.1672	−0.0452	−0.1796	0.3130	0.2732	0.0835
70	0.0963	0.0439	−0.0815	0.2298	0.1982	0.0430
80	0.0448	0.0789	0.0063	0.1471	0.1303	0.0194
90	0.0128	0.0576	0.0486	0.0692	0.0670	0.0072
100	0.0002	0.0051	0.0097	0.0052	0.0061	0.0006
eq.(*)	(32)	(19)	(25)	(22)	(31)	(28)

* Number of equation used to compute.

Table 5 Selected Spectral Density: Trend Stationary Case

Freq-uency	With Intercept		No Intercept		
	OLS	CF	OLS	CF (q=0)	CF (q=n)
0.0000	0.0000	−0.0000	1.0100	0.8547	30.8383
0.0025	0.0083	2.6398	0.8234	3.5655	26.9042
0.0050	0.1076	7.2958	0.5077	8.1530	17.7242
0.0075	0.3790	8.4663	0.4602	8.8433	8.7401
0.0100	0.7223	5.2728	0.7151	5.1390	3.6033
0.0125	0.9378	1.7033	0.9630	1.5374	2.1142
0.0150	0.9663	0.8743	1.0012	0.9830	2.0216
0.0175	0.9311	1.7663	0.9364	1.9241	1.7850
0.0200	0.9458	1.9737	0.9364	1.9237	1.3507
0.0225	0.9906	1.2320	0.9920	1.1162	1.1210

Table 6 Selected Autocorrelations: Trend Stationary Case

	With Intercept		No Intercept		
Lag	OLS	CF	OLS	CF (q=0)	CF (q=n)
1	−0.0198	0.1275	−0.0097	0.1410	0.3952
10	−0.0162	0.0903	−0.0070	0.1024	0.3415
20	−0.0124	0.0507	−0.0042	0.0612	0.2830
30	−0.0087	0.0145	−0.0016	0.0234	0.2269
40	−0.0055	−0.0165	0.0006	−0.0093	0.1743
50	−0.0027	−0.0405	0.0024	−0.0352	0.1265
60	−0.0004	−0.0559	0.0036	−0.0527	0.0846
70	0.0011	−0.0608	0.0041	−0.0601	0.0498
80	0.0018	−0.0536	0.0038	−0.0556	0.0233
90	0.0015	−0.0325	0.0026	−0.0375	0.0063
100	0.0002	0.0043	0.0003	−0.0042	0.0000
eq.(*)	(34)	(36)	(35)	(38)	(37)

* Number of equation used to compute.

Fig. 1　Trend Removal by Models with an Intercept: Random Walk

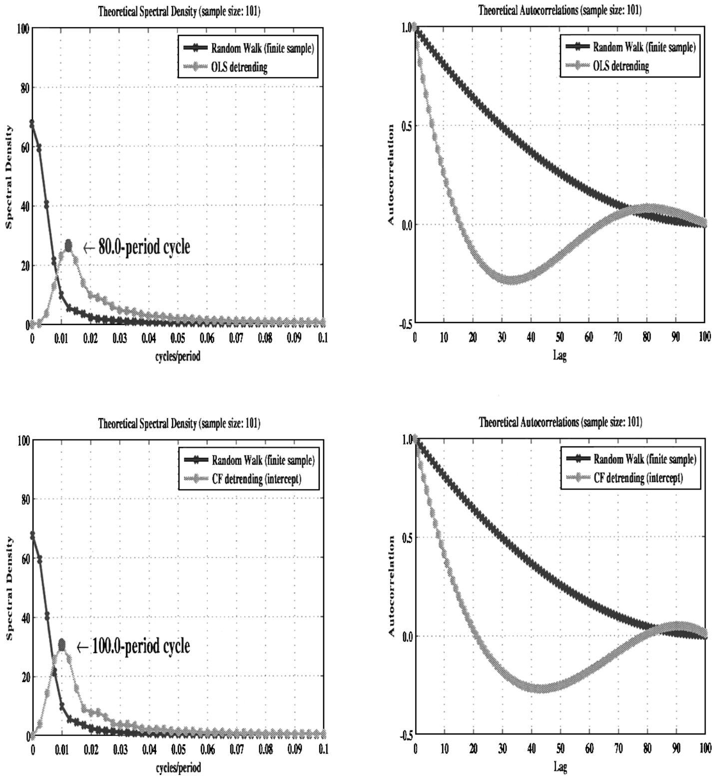

Fig. 2 Trend Removal by Models with No Intercept: Random Walk

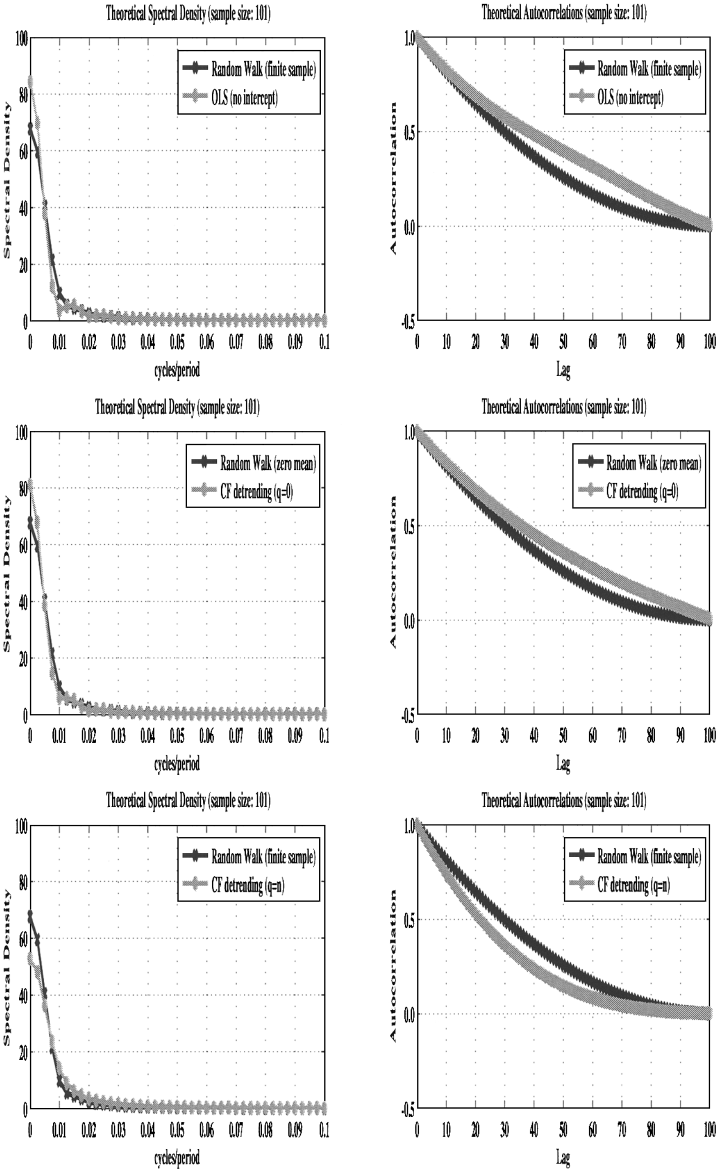

Fig. 3　Trend Removal by Models with an Intercept: Trend Stationary

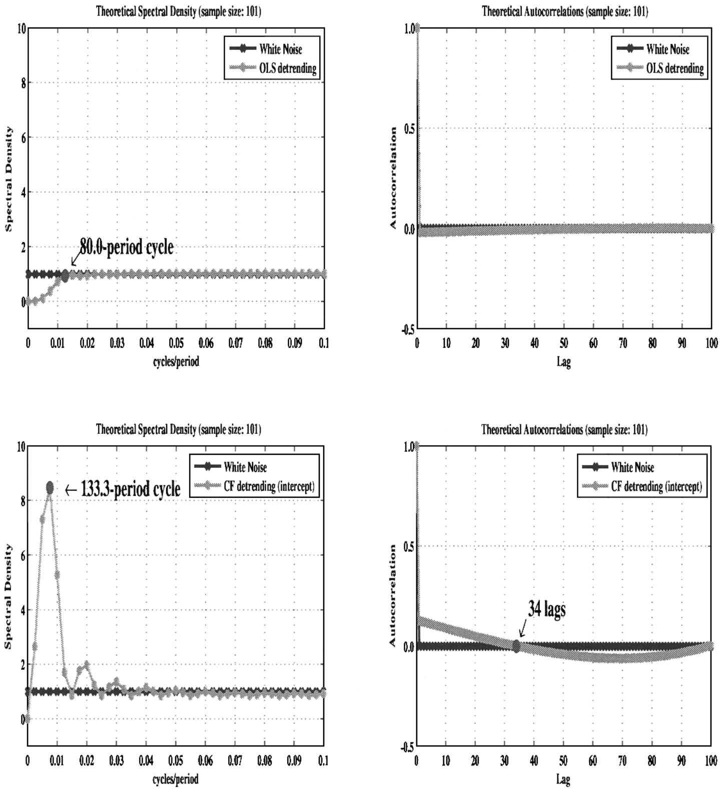

Fig. 4 Trend Removal by Models with No Intercept: Trend Stationary