Linear Data Transformations Used in Economics

Darrel Cohen

Federal Reserve Board

December 2001

Abstract: This paper examines the properties of standard data transformations—such as growth rates and moving averages—used by applied economists. Because many resources are devoted to understanding the economic significance of incoming data by government and financial-market economists, for example, this paper considers data filters that do not drop recent observations, in contrast to the approximately "ideal" measures recently developed in the literature. Using frequency-domain techniques, it is established that moving averages of multi-period growth rates can attenuate the bias and phase shifts introduced by common data filters.

Correspondence information should be sent to Darrel Cohen, Stop 83, Division of Research and Statistics, Federal Reserve Board, Washington, D.C. 20551. Telephone: 202-452-2376; Fax: 202-452-5296; E-mail: dcohen@frb.gov.

The views expressed in this paper are those of the author and do not necessarily represent those of the Board of Governors or the staff of the Federal Reserve System.

Linear Data Transformations Used in Economics

I. Introduction

Economists use many transformations of time series data to help extract economically relevant information. Such transformations include 4-quarter and 1-quarter growth rates, 6-month and 3-month moving averages, trailing and centered moving averages, moving averages of growth rates, and so on.

Ultimately, the best data transformation or filter is the one that helps the most at addressing the economic question of interest. For example, when considering the cyclical properties of the real economy, filters that highlight the cyclical features of the data are desirable. On the other hand, when considering whether price inflation is changing because of cyclical or longer-term factors, it is desirable to apply filters that bring out the cyclical and trend-like features of the data, together or separately. Of course, choosing a good filter should not be viewed as the only step in determining the economic significance of data; rather it can complement economic insight and knowledge of other factors, such as how the data are constructed.

In the recent literature, the frequency-domain theory of linear filters has been used to develop approximations to "ideal" data transformations or filters. Intuitively, ideal filters pass through information without distortion from the frequencies that are of interest to the analyst (such as business-cycle frequencies) and completely block out all other information. Indeed, the recently-developed filters have deepened our understanding of meaningful historical economic relationships, discussed below.

However, the new filters are of very limited use for economic agents such as policymakers and financial-market participants, who devote many resources to understanding the significance of incoming economic data, because the filters either completely drop recent observations or generate unreliable estimates of them. This paper examines filters subject to the constraint that many economic analysts will continue to utilize the latest published data, relying on less-than-ideal data filters. It lays out the properties of standard data transformations and suggests improvements using frequency-domain techniques. One property is the degree of bias, that is, the amount by which the

filter employed deviates from the "ideal" filter on average over the frequency range of interest; to our knowledge, the concept of bias in the frequency domain has not been examined before. In addition, the timing or phase shift between a raw data series and its filtered value is explored. The key finding of this paper is that moving averages of multi-period growth rates can attenuate the bias and phase shifts introduced by common data filters.

II. A Brief History of Frequency

The idea that even simple transformations can change the properties of the underlying data series in surprising and undesirable ways goes back to Slutzky's work in the 1930s. His work implies, for example, that taking the 2-quarter moving average of the 1-quarter growth rate of a white-noise quarterly series, ϵ_t, creates a new series with a seasonal pattern, and taking a 4-quarter growth rate of ϵ_t creates a new series with a business cycle pattern but no seasonal pattern.

The message is that standard transformations may introduce "spurious" cycles in economic data and thus possibly generate misleading economic interpretations. However, if the underlying series has pronounced cyclical and seasonal components, then data transformations need not introduce spurious patterns. The techniques of spectral analysis discussed below are useful precisely because they allow us to determine whether cyclical or seasonal patterns (or for that matter any other patterns of interest) in the data could have been introduced inadvertently through a seemingly innocuous data transformation.

Recent papers explore the possibility of approximating an *ideal* filter. Indeed, papers by Baxter and King (1999) and by Christiano and Fitzgerald (1999) provide finite-sample approximations to an ideal filter. Both Baxter-King (BK) and Christiano-Fitzgerald (CF) focus on ideal business-cycle filters and show that certain important economic relationships, like the negative Phillips curve correlation between inflation and unemployment rates, show up strongly at business cycle frequencies–i.e., when the "ideal" business-cycle filter has been applied to the raw inflation and unemployment rate data–but not at lower frequencies.[1]

---

1. It is arguable that the strong negative correlation at the business cycle frequencies and non-negative correlation at lower frequencies reflect the idea that inflation expectations are dominated by slow-

Although the exact specifications of the ideal filters differ in the two papers for technical reasons, both are two-sided moving averages. With quarterly data, the BK ideal business-cycle filter is a weighted sum of the current value and 12 leads and lags of the time series with symmetric weights. Through extensive experimentation with quarterly macroeconomic data, it was determined that 12 was the value that produced filtered series that best satisfied the BK optimization criteria. A major disadvantage of the BK filter to analysts who focus on recent observations of a time series is that three years worth of data at the ends of the sample period are dropped in computing the filter.

For most macroeconomic time series, the CF ideal business-cycle filter is a weighted sum of the current and *all* leading and lagged values of the data. In contrast to the BK filter, the CF filter in these cases is asymmetric, discarding no data at the ends of the sample period. However, CF note that the filtered values of the last couple of observations of the y_ts are not reliably estimated; put another way, their method yields reliable filtered values for y_{T-2} and earlier observations. Examples in the CF paper using monthly, quarterly, or annual data each take y_{T-2} as the observation 2 years from the end of the sample; although the filtered value of y_{T-2} makes use of all data subsequent to period T-2, it is implied that filtered values subsequent to that period are not very reliable.

In addition, the recent research establishes that the famous Hodrick-Prescott (HP) filter–which also is a business cycle filter and does not discard data at the ends of the sample period–is not "ideal" but, as a practical matter, appears to be a good approximation with quarterly macroeconomic data. In addition, the HP filter (with parameter $\lambda = 1600$) appears to be a reasonable approximation to the ideal filter that cuts off all information from periods longer than 32 quarters and allows through all information from periods shorter than 32 quarters (a so-called "high-pass" filter). However, the HP filter is not without problems: its filtered estimates become increasingly unreliable at the end of the sample period and it does not appear appropriate to use with annual data.

---

moving components.

III. A Brief Overview of Spectral Analysis

Spectral analysis provides the framework for addressing the key issues of this paper. At its most basic level, spectral analysis is based on a fundamental theorem that any stationary time series, y_t, loosely speaking, can be represented as the sum of (a possibly uncountably infinite number of) uncorrelated cosine and sine waves of different frequencies and amplitudes.[2] A formal statement and proof of the representation theorem can be found in Priestly (1981, chapter 4), for example. Because of this result, the variance of a stationary time series is related to the amplitudes–that is, the maximum heights and depths–of the cosine (or sine) waves of all frequencies or equivalently, of all periods.

Low frequencies capture data patterns that repeat only over long periods (i.e., that capture trend or slow moving components of the data); business-cycle frequencies capture data patterns or cycles that repeat roughly once every 6 to 32 quarters or 18 to 96 months (values used in the recent literature and attributed by Baxter and King (1999) to the NBER research of Arthur Burns and Wesley Mitchell); and seasonal frequencies capture cycles that repeat within a year.

The contributions that cycles of each frequency make to the total variance of a time series is described by the *spectrum*, a curve that plots contributions to variance on the vertical axis against frequency, measured from lowest to highest (or, equivalently, period from longest to shortest), on the horizontal axis.[3] In what follows, ω denotes frequency, measured in radians per unit time; thus $2\pi/\omega$ is the period, or length of time required for a sinusoidal process to repeat a full cycle. The spectrum can take many shapes. For example,

---

2. In a simple case, $y_t = A_1 \cos(\omega_1 t + \theta_1) + A_2 \cos(\omega_2 t + \theta_2) + A_3 \cos(\omega_3 t + \theta_3)$, that is, y_t is the sum of three cosine functions of different frequencies (ω_i) and phases (θ_i), with each function multiplied by A_i (the amplitude). If $\theta_1 = \theta_2$, $\omega_1 = \omega_2$, $\omega_3 = \theta_3 = 0$, and $A_1 = -A_2$ then y_t is a horizontal line. Chatfield (1996) points out that in general, if the A_i's and θ_i's are constants, $E(y_t)$ will change with time, implying that y_t is not stationary. More fundamentally, y_t is not even a random process unless the A_i's or θ_i's as random variables; further, the A_i's or θ_i's must have properties that ensure the stationarity of y_t.

3. The spectrum technically is defined over frequencies ranging from $-\pi$ to π (radians per unit time), and the area under the spectrum is the total variance of the time series. Because the spectrum is symmetric about the zero frequency, it normally is plotted only for frequencies from 0 to π, or equivalently it is plotted for periods from infinity to 2 periods.

if the spectrum is a horizontal line, then fluctuations at each frequency contribute the same amount to the overall variance of a series; in this case the series is *white noise* (just as the color white is made up of equal contributions from all colors). Granger (1966) argues that the "typical spectral shape" of an economic series in level form, such as GDP, is very high at low frequencies, falls steeply as frequency is increased, and eventually flattens out; such a sample spectrum either means that the variance of the series is dominated by low frequency (trend-like) cycles or that the series is non-stationary.

IV. Data Transformations or Filters

Data transformations or *filters* modify the spectrum of the underlying series, multiplying that spectrum by a factor greater than one at some frequencies and less than one at others. If the filter is *linear* (i.e., if the filtered series is a linear combination of elements of the original series), then it is possible to determine analytically how much the underlying theoretical spectrum is distorted at each frequency.

An *ideal* filter only passes through the frequencies that are of interest to the analyst, such as business-cycle frequencies, and completely blocks out all other information. Moreover, within the frequency band of interest, the ideal filter does not re-weight the underlying information from each frequency. Chart 1 illustrates an ideal *business-cycle* filter function for quarterly data (top panel) and monthly data (bottom panel). Owing to limitations of our plotting software, the horizontal axis of the figures in Chart 1 (as well as other figures in Charts 2 and 3) use "PI" to denote π. The horizontal axis measures frequency, ω, from 0 to π (radians per quarter with quarterly data or radians per month with monthly data) or equivalently measures period from infinity to 2 (quarters or months, depending on the data).

In each case shown, the ideal filter function equals one inside the range of business-cycle frequencies and zero outside the range. Of course, many other ideal filters with this one-zero property are possible: these are referred to as ideal "low-pass" filters (that highlight only the slow-moving components of the data), ideal "high-pass" filters (that highlight only the high-frequency components of the data), and ideal "band-pass" filters (of which the ideal business-cycle filter is a special case). As shown below, the filter functions

associated with standard data transformations are far from "ideal" but, with slight modifications, can be made to much better approximate an ideal filter.

V. Bias in the Frequency Domain

We now introduce the concept of *bias* in the frequency domain.

Definition: Denote the average value of the filter function over the frequency range of interest as AvgF. The bias of the filter function is AvgF - 1.

That is, bias is the difference between the average value of the filter function and one (the value of the ideal filter) over the frequency range of interest. It follows almost trivially that, over the entire frequency range from 0 to π (radians per unit time), the ideal low-pass, high-pass, and band-pass filter functions are biased (because each has a value of one only over a subset of the entire frequency range). A more important result, presented in the following proposition, is that the direction of bias determines whether the data transformation leads to a more or less variable time series; as we will see below, growth-rate filters have positive bias and, hence, are variance increasing, while moving average filters have negative bias and are variance decreasing.

Proposition 1: The variance of the data series generated by a data transformation exceeds that of the underlying stationary series if AvgF > 1 (over the entire range of frequencies) and is less than that of the underlying series if AvgF < 1.

   Proof: Recall that the area under the spectrum is the variance of a series and that the spectrum of a transformed data series in general exceeds the spectrum of the underlying series for some frequencies and is less at others; put another way, the filter function of the transformed series increases the spectral values of the underlying series at some frequencies and lowers them at others. On balance, the increases dominate, and variance is increased, if AvgF exceeds one (over the entire range of frequencies); similarly, the decreases dominate if AvgF is less than one.

VI. Growth Rate Filters

Consider a typical data transformation of the level of a quarterly series, like GDP, into growth rates. Let y_t denote the logarithm of the variable; thus $x_t = y_t - y_{t-1}$ is approximately the non-annualized growth rate of the variable. Here we have used a first-difference filter,

which is linear. There are two points to emphasize.

First, the annualized growth rate is given by $(1 + x_t)^4 - 1 = 4x_t + 6x_t^2 + 4x_t^3 + x_t^4$ and thus is not a linear filter; strictly speaking, the theory of linear filters does not apply to the annualized growth rate, but we will consider the linear part—namely, $4x_t$, below. Second, at each frequency, ω, the spectrum of the original level series is multiplied by the factor, $2 - 2 \cos\omega$, for the non-annualized growth rate and by the factor, $16 (2 - 2\cos\omega)$, for the linear part of the annualized growth rate (as proven below). These are called "filter functions," and several are shown in the accompanying charts.

By plugging in values for ω in the case of the nonannualized growth rate, it is seen that the spectrum of x_t (the growth rate) is derived by multiplying the spectrum of y_t (the log level) by factors that grow monotonically from 0 to 4 as frequency rises from 0 to π; for the linear part of the annualized growth rate, the multiplicative factors grow from 0 to 64 (see the dotted line in the top left panel of chart 2). The contribution to overall variance from the zero frequency is completely "killed" or eliminated by the data transformation; this is desirable because transforming levels data into growth rates is typically done to eliminate long-term stochastic trends or unit roots.[4]

The *cost* of this data transformation is that contributions from certain business-cycle frequencies also are attenuated, while contributions from other business-cycle frequencies and high frequency variation are greatly magnified. Indeed, with quarterly data the multiplicative factors for the linear part of the annualized growth rate are less than the "ideal" value of one for periods longer than about 25 quarters (or frequencies less than .25) and greater than one for periods shorter than 25 quarters. The dividing line between attenuation and magnification arising from this simple growth-rate transformation appears quite arbitrary.

---

4. Strictly speaking, contributions from the zero frequency are completely eliminated only if the original series is stationary. For example, if the original series is I(2), then first differencing produces a non-stationary I(1) series whose spectrum is infinite at the zero frequency. Also as a practical matter, standard methods of estimating the spectrum do not produce a zero spectral value at the zero frequency even for a stationary series because the estimate at the zero frequency is a weighted average of estimated spectral values near and including zero.

These results readily generalize, as summarized in the following proposition. In the subsequent discussion, the operator, L, is defined by $L^k z_t = z_{t-k}$ for both positive and negative integer values of k.

Proposition 2: The spectrum of the k-period nonannualized growth rate, $x_t = y_t - y_{t-k} = (1 - L^k)y_t = c(L)y_t$, is the spectrum of y_t multiplied by the factor $2 - 2\cos k\omega$.

Proof: If $s_z(\omega)$ denotes the spectrum of z, then, as established in several textbooks, such as Hamilton (1994, chapter 6), $s_x(\omega) = c(e^{i\omega})c(e^{-i\omega}) s_y(\omega)$, where $c(e^{i\omega})c(e^{-i\omega})$ is the filter function or the squared gain. In our case, $c(e^{i\omega}) = 1 - e^{ik\omega} = 1 - \cos(k\omega) - i\sin(k\omega)$, from which it follows that $c(e^{i\omega}) c(e^{-i\omega}) = 2 - 2\cos k\omega$.

This sort of filter is often employed in data presentations. For example, with quarterly data and $k = 4$, $y_t - y_{t-k}$ is approximately the 4-quarter percent change; with monthly data and $k = 12$, $y_t - y_{t-k}$ is approximately the twelve-month percent change.

Proposition 2 allows us to compute the average value of the filter function of the k-period nonannualized growth rate. For all k, the average value is 2 over the entire frequency range, and the average value of the 4-quarter growth rate over the quarterly business-cycle frequency range of $\pi/16$ to $\pi/3$ radians per quarter is 2.9.[5] Moreover, the average value also exceeds the "ideal" value of 1 for all values of k over the quarterly business-cycle frequencies. However, over the low-frequency range of 0 to $\pi/16$ radians per quarter, the average value is 0.2 for $k = 4$.[6] Because the average values generally differ from the "ideal" value of 1, the growth-rate filters over the frequency ranges commonly of interest are biased: the magnification of the spectrum of the underlying series at some frequencies is not offset on average by the attenuation at others. We summarize this discussion in propositions 3 and 4:

Proposition 3: Over the frequency ranges commonly of interest, the filter function of the k-period nonannualized growth rate is biased. Over the business-cycle frequencies, the average value of the filter function exceeds 1 for all values of k.

---

5. The average value of the former is: $(1/\pi) \int_0^\pi (2 - 2\cos k\omega)d\omega = 2$ for $k = 1,2,3,...$. The average value of the latter is: $[(\pi/3 - \pi/16)]^{-1} \int_{\pi/16}^{\pi/3} (2 - 2\cos 4\omega)d\omega = 2.9$.

6. For $k = 10$, the average value of the filter function is about 1 over the low frequencies.

Proposition 4: Because the average value of the filter function of the k-period nonannualized growth rate is 2 over the entire frequency range, the variance of the series generated by a k-period nonannualized growth rate filter always exceeds the variance of the underlying series (assuming stationarity).[7]

Further, as shown in the top two panels of chart 2, the filter functions for the cases k = 4 (used with quarterly data) and k = 12 (used with monthly data) have multiple peaks (at a value of 4) and "kill" the contribution to overall variance at the zero and seasonal frequencies; because of the latter, they sometimes are called seasonal difference filters.[8] As noted in Granger and Newbold (1986), it is sometimes argued that all this killing (i.e., giving a zero weight to contributions from the seasonal frequencies) is not necessarily a good thing, because even a white noise series has some of its variance explained by variation at the seasonal frequencies.

Also, as seen by careful inspection of the top left panel of chart 2, the 4-quarter percent change yields a smoother transformed series relative to the *annualized* 1-quarter percent change series, because the value of the 4-quarter percent change filter function (read off the left axis) is less than the value of the annualized 1-quarter growth filter function (read off the right axis) at every frequency except zero. Similarly, the 12-month percent change smooths relative to the annualized 3-month percent changes, as seen from the top right panel of chart 2.

The 4-quarter percent change not only smooths relative to that of the linear part of the annualized 1-quarter percent change, but it also produces a series more in time synchronization (that is, more in *phase*) with the underlying series. In general, growth-rate filters generate a time series that *leads* the underlying series (i.e., both the peaks and troughs come earlier), albeit by a lesser amount as the value of k increases; these ideas are

---

7. For example, if the underlying series is white noise with variance σ^2 (and constant spectral value of $\sigma^2/2\pi$), then the variance of the transformed series is $(\sigma^2/2\pi) \int_{-\pi}^{\pi} (2 - 2\cos k\omega)d\omega = 2\sigma^2$.

8. The seasonal frequencies, ω_s, are defined as $2\pi k/p$ where k = 1,2,3, ... and p is the period of the seasonal, taking the value of 4 for quarterly data and 12 for monthly data. For example, with quarterly data the seasonal frequencies are $\pi/2$ and π radians per quarter, corresponding to periods $(2\pi/\omega_s)$ of 4 quarters and 2 quarters, respectively. With monthly data, there are 6 seasonal frequencies, starting with $\pi/6$ (or equivalently, a period of 12 months).

implications of the next proposition (where variables are again expressed in logs).[9]

Proposition 5: Let $\theta_k(\omega)$ be the phase spectrum of $\lambda(y_t - y_{t-k})$, where λ is a constant; then $\theta_k(\omega) = \tan^{-1}[\sin(\omega k)/(1-\cos(\omega k))] > 0$ for $0 < \omega < \pi/k$. Also, $d\theta_k(\omega)/dk < 0$ for $\omega > 0$.

Proof: Let $x_t = \lambda y_t - \lambda y_{t-k} = h_0 y_t + h_1 y_{t-k}$, where $h_0 = \lambda$ and $h_1 = -\lambda$ and $h_j = 0$, otherwise. The frequency response function is: $H(\omega) = \lambda e^{-i\omega(0)} - \lambda e^{-i\omega(k)} = \lambda(1 - \cos\omega k + i\sin\omega k)$. By definition, the phase is thus $\theta_k(\omega) = \tan^{-1}[\sin(\omega k)/(1-\cos(\omega k))]$. Since $1-\cos(\omega k) > 0$ (except for $\omega = 0$) and $\sin(\omega k) > 0$ for $0 < \omega < \pi/k$, it follows that $\theta_k(\omega)$ is positive (over the interval defining the principal branch of the inverse tangent function), implying that the filtered data lead the underlying data for $0 < \omega < \pi/k$.

Let $u(k) = \sin(\omega k)/(1-\cos(\omega k))$. Thus, $\theta_k(\omega) = \tan^{-1} u(k)$. From the calculus of trigonometric functions, we know that $d\theta_k(\omega)/dk = (1/(1+u^2))du/dk$. Carrying out the differentiation, we get that $d\theta_k(\omega)/dk = -\omega/2 < 0$ (for $\omega > 0$).

Intuitively, the reason that growth rates typically lead is that, as the underlying series approaches a peak, the positive increments to the series diminish in value (the second derivative is negative in the neighborhood of a maximum); hence, the increments drop off before the actual series.

In summary, these growth rate filters have the desirable property of removing very long run stochastic trends (or attenuating them if the underlying series is nonstationary). Further, the smoothing and better time synchronization achieved by the 4-quarter growth rate filter *relative* to the annualized 1-quarter growth rate filter (or the 12-month growth rate filter relative to the annualized 3-month growth rate filter) often is desirable. However, the costs of all growth-rate filters are the introduction of bias (in the sense that the magnification of the spectrum of the underlying series at some frequencies is not offset on average by the attenuation at others) and timing shifts (that is, the transformed series leads the underlying series).

---

9. The so-called phase spectrum captures the degree of time asynchronization between a series and the filtered value of it. In general the phase varies with frequency, making it a tricky concept; for example the business cycle frequency components of a filtered series may lead the underlying series but higher frequency components may lag it. To compute the phase, $\theta(\omega)$, for the linear system $x_t = \Sigma h_j y_{t-j}$ (with summation over j running from $-\infty$ to ∞), start by taking the Fourier transform of the impulse response function $\{h_j\}$. This is known as the frequency response function, $H(\omega)$, and is generally a complex variable, $R(\omega) + iI(\omega)$. The phase, $\theta(\omega)$, is defined as $\tan^{-1}[I(\omega)/R(\omega)]$.

VI. Moving Averages

Intuitively, arithmetic moving averages are used to highlight underlying trend and smooth out local fluctuations; indeed in some textbooks, the moving average is called a smoothing filter. The following three propositions and corollary summarize important points about moving averages. Both trailing and centered moving averages with uniform (constant) weights are evaluated.

Proposition 6: The spectrum of an n-term moving average $[x_t = (y_t + y_{t-1} + ... + y_{t-(n-1)})/n]$ is derived by multiplying the spectrum of the original series (y_t) by the factor:
$[n + 2\cos(n-1)\omega + 4\cos(n-2)\omega + 6\cos(n-3)\omega + 8\cos(n-4)\omega + ... + 2(n-1)\cos\omega]/n^2$.

Proof: Consider $x_t = (y_t + y_{t-1} + ... + y_{t-(n-1)})/n = (1/n)(1 + L + L^2 + ... + L^{n-1})y_t = m(L)y_t$. Thus the spectrum of the n-term moving average, $s_x(\omega)$ is given by $m(e^{i\omega}) m(e^{-i\omega}) s_y(\omega)$, where $m(e^{i\omega}) m(e^{-i\omega})$ is the filter function of the moving average. We start with the case of n = 2. In this case, $m(e^{i\omega}) m(e^{-i\omega}) = (1/4)[1 + e^{i\omega}][1 + e^{-i\omega}] = (1/4)(2 + e^{i\omega} + e^{-i\omega}) = (1/4)(2 + 2\cos\omega)$. Subsequent steps, not shown, are a straightforward application of proof by induction.

Proposition 6 implies that the variance of a time series is reduced by applying a trailing moving average to it (that is, the resulting transformed series looks smoother than the original series). This occurs because the moving-average filter downweights the contribution to overall variance from high frequencies *relative* to those from low frequencies; and, in addition to generally declining with frequency, the moving-average filter function starts with a value of unity at the zero frequency, implying that the average value of this filter function is always less than one.[10] Generally, the more terms present in the moving average, the more rapidly the filter function declines with frequency and, hence, the smoother the resulting time series. These claims are proved in the following corollary.

Corollary: The value of the filter function of the n-period trailing moving average is less than 1 at all frequencies, except 0, and the average value is 1/n. As n rises, the filter function declines more rapidly with frequency, when evaluated at low frequencies.

---

10. If the underlying series is white noise with variance σ^2, then the variance of the n-term moving average is determined as the product of $\sigma^2/2\pi$ and the integral from $-\pi$ to π of the above filter function; the resulting product is σ^2/n. Thus, even a 2-term moving average cuts the variance of an underlying white noise series in half. If the underlying series is a stationary first-order autoregressive process with parameter ρ (and $|\rho| < 1$), then the variance of the 2-term moving average is $(1 + \rho)/2$ times the variance of the underlying series.

Proof: By direct substitution, we see that the value of the filter function of the n-period moving average (as specified in proposition 6) at $\omega = 0$ is unity. This is the maximum value of the filter function over all frequencies from 0 to π, because the maximum value of $\cos(x)$ is 1 (at $x = 0$) and in general the cosine terms in the filter function are thus less than 1. Hence, the average value of the filter function is less than 1. The average value is $1/n$, derived by integration of the filter function from 0 to π (and division by π).

Finally, differentiate the filter function with respect to ω. Evaluated at $\omega = 0$, the first derivative is zero and the second derivative is negative, implying that the filter function attains a maximum at $\omega = 0$. Further, differentiation of the second derivative (evaluated at $\omega = 0$) with respect to n yields a negative number, implying that the filter function falls off more rapidly as n rises.

Examples of 3-, 4-, and 6-period trailing moving average filter functions are presented in the lower left panel of chart 2. These examples suggest that for quarterly data, a 4-quarter moving average is sufficient to bring out the underlying trend and minimize the contribution from higher frequencies, because this filter generally gives very little weight to contributions from frequencies above the business-cycle cutoff value of $\pi/3$ and completely eliminates contributions from the quarterly seasonal frequencies; for monthly data, at least a 6-month moving average of the levels data is appropriate.

The second key property of moving averages, summarized in proposition 7, is that the spectrum of a trailing moving average of a stationary series is the same as the spectrum of a centered or two-sided moving average with equal (symmetric) weights, ignoring important issues that arise at the end points of a time series; thus both moving averages produce the same amount of smoothing.

Proposition 7: The filter functions of an n-period trailing moving average and an n-period centered moving average with equal weights are identical (ignoring end-point issues).

Proof: Assume that n is an odd integer. We know the filter function of the trailing moving average from proposition 6; the filter function of the centered moving average also is established by induction. We will show the equivalence for the $n = 3$ case. In this case, the centered moving average is given by: $(y_{t+1} + y_t + y_{t-1})/3 = (1/3)(1 + L + L^{-1})y_t = m(L)y_t$. Thus, $m(e^{i\omega}) = (1/3)(1 + e^{i\omega} + e^{-i\omega})$, and the filter function becomes $m(e^{i\omega})m(e^{-i\omega}) = (1/9)(1 + 4\cos\omega + 4\cos^2\omega)$. However, using the identity $\cos2x = 2\cos^2x - 1$, this can be re-written as $(1/9)(3 + 4\cos\omega + 2\cos2\omega)$, completing the proof.

The final key property of moving averages, summarized in proposition 8, is that

centered moving averages (with symmetric, but not necessarily uniform, weights) generally are in perfect time synchronization with the underlying economic time series, and trailing moving averages generally lag it.[11] Moreover, the lag increases with the number of terms in the trailing moving average.

Proposition 8:   Centered moving averages with symmetric weights cannot lag, and generally are in perfect time synchronization with, the underlying series, but trailing moving averages with uniform weights generally lag it.

   Proof: As shown in Granger and Newbold (1986, chapter 2), the tangent of the phase of an n-term *trailing* moving average with uniform weights is [-Σsinjω/Σcosjω], where the summation over j is from 0 to n-1; the phase is typically, but not always, negative, implying that the filtered series generally lags the underlying series.
   We now establish that the tangent of the phase of an n-term *centered* moving average is 0, implying that the phase itself is either zero (implying perfect time synchronization) or π (implying that the filtered series leads the underlying series), depending on the frequency. We assume that the moving average has uniform weights, but it will become clear that the results hold for symmetric (but not necessarily uniform) weights as well. (Recall that Baxter and King (1999) develop a centered moving-average filter with symmetric weights.)
   Let $x_t = \Sigma h_j y_{t-j}$ where the summation is over j from -n to n (a total of 2n + 1 terms) and where $h_j = 1/(2n+1)$ for all j. The frequency response function is: $H(\omega) = \Sigma h_j e^{-i\omega j} = (2n+1)^{-1}[e^{i\omega n} + e^{i\omega(n-1)} + ... + e^{-i\omega 0} + e^{-i\omega} + ... + e^{-i\omega n}] = (2n+1)^{-1}[1 + 2\cos \omega n + 2\cos\omega(n-1) + .. + 2\cos\omega]$. The final equality follows from the property that $\xi(e^{iz} + e^{-iz}) = 2\xi\cos z$, where ξ is a constant; the complex components cancel provided that e^{iz} and e^{-iz} are each multiplied by the same factor. This is why the results carry over to the case of symmetric weights. We thus have established that the frequency response function is real and therefore the tangent of the phase spectrum is zero, implying that the phase itself is 0 or π.

VII. Combining Differences and Moving Averages

Combining differences and moving averages to transform the data generates filter functions that rapidly become very complicated as the order of differencing increases, although the combined filter function is always the product of the two separate ones. The following proposition presents the combined filter.

---

11. Examples of atypical cases are shown in the appendix chart. Both panels plot an underlying and a filtered series against time on the horizontal axis. The top panel illustrates that a 3-term centered moving average of a choppy time series (specified as $10\cos[(3/4)\pi t]$) can lead the underlying series rather than be in perfect synchronization; Chatfield (1996, chapter 9) derives the properties of the phase of a 3-term centered moving average. The bottom panel illustrates that a 4-term trailing moving average can be in perfect synchronization.

Proposition 9: The spectrum of the n-term trailing moving average of the k-period nonannualized growth rate is the spectrum of the original series (y_t) multiplied by the factor: $(2 - 2\cos k\omega)[n + 2\cos(n-1)\omega + 4\cos(n-2)\omega + 6\cos(n-3)\omega + ... + 2(n-1)\cos\omega]/n^2$.

Proof: Let $x_t = [(y_t - y_{t-k}) + (y_{t-1} - y_{t-k-1}) + ... + (y_{t-(n-1)} - y_{t-k-(n-1)})]/n$

$$= [1 - L^k + L - L^{k+1} + ... + L^{n-1} - L^{k+(n-1)}] (1/n) y_t$$

$$= [1 - L^k] [1 + L + ... + L^{n-1}] (1/n) y_t,$$

which implies, using propositions 2 and 6, that the filter function of x_t (the n-term trailing moving average of the k-period growth rate) factors into the product of the separate filter functions of the n-term trailing moving average and the k-period nonannualized growth rate.

Note that the combined filter is zero at the zero frequency (owing to the differencing component of the filter), implying that it renders stationary an I(1) stochastic process. Baxter and King (1999) consider the ability of a filter to result in a stationary series an important property. In addition, the combined filter simplifies greatly in the case of first differences:

Corollary: The filter function of the n-period moving average of the nonannualized first-difference is given by $(1/n^2)(2 - 2\cos n\omega)$. The average value of this filter over all frequencies (from 0 to π) is $2/n^2$.

Proof: Because of the fact that $(1/n)(1+L+L^2+...+L^{n-1})(1 - L) = (1/n)(1 - L^n)$, the filter function of the n-period moving average of the first difference is proportional to the filter function of the n-period difference, whose form we know from proposition 2. Integration of the filter function over all frequencies establishes that the average value is $2/n^2$.

Examples of combined filters relevant to monthly data are shown in the bottom right panel of chart 2, and examples relevant to quarterly data are shown at the top panel of chart 3. The solid line in the bottom right panel of chart 2 is the filter function for a 3-month moving average of the 12-month percent change, a measure often emphasized by price analysts on Wall Street. This combined filter places much less weight on high-frequency variations than does the 12-month percent change but retains much of the same distortion over the business cycle and lower frequency range; as discussed below, other combined filters have more desirable properties.

As noted above, the average value of the filter function of the k-period growth rate

exceeds one over the entire range of frequencies as well as over the frequencies normally of most interest (i.e., business-cycle and lower frequencies), while the average value of the filter function of moving averages is less than one. Thus, combined filters involving k-period growth rates with k > 1 offer the possibility of having mean values near one, thereby not badly distorting the spectral information of the underlying series *on average*.[12] This is one of our key insights and is summarized in proposition 10; the proposition will be proved assuming quarterly time series data but also holds for monthly data as well (with examples given below).

Proposition 10: For k > 1 and over the business-cycle and lower frequency range, there exists integer values for k and n such that the average value of the filter function of the combined n-term moving average of the k-period nonannualized growth rate is near one.

 Proof: As noted above, we will assume quarterly data, in which case the frequency range of interest runs from 0 to $\pi/3$ (with a corresponding period of at least 6 quarters). We will show that the average value of the filter function is exactly unity in the $n = k = 2$ case, and that the average value is approximately unity for other (n,k) pairs as well.

 For the $n = k = 2$ case, the average value of the filter function in proposition 9 is:

$$(3/4\pi) \int_0^{\pi/3} (2 - 2\cos2\omega)(2 + 2\cos\omega)\, d\omega = 1, \text{ using } \cos x \cos y = [\cos(x+y) + \cos(x-y)]/2.$$

By integrating the combined filter function and using the same trigonometric identity, we can find other (n,k) pairs that yield an average value near 1 for $0 \le \omega \le \pi/3$. Because 4-quarter percent changes are often used in applied work, it is interesting to consider the case of k = 4. For n = 4, the average value of the filter function is:

$$(3/16\pi) \int_0^{\pi/3} (2 - 2\cos4\omega)(4 + 2\cos3\omega + 4\cos2\omega + 6\cos\omega)\, d\omega = 1.27,$$

and the average value for n = 5 is:

$$(3/25\pi) \int_0^{\pi/3} (2 - 2\cos4\omega)(5 + 2\cos4\omega + 4\cos3\omega + 6\cos2\omega + 8\cos\omega)\, d\omega = 0.87.$$

Thus, for k = 4, either a 4-term or 5-term moving average yields an average value of the filter function that is near unity.

---

12. Using the above corollary for the k = 1 case, the combined filter has an average value of $2/n^2$ over the entire range of frequencies, a number less than one.

The numerical examples presented in chart 2 suggest that to minimize the average distortion over the frequencies normally of most interest with monthly data–i.e., frequencies less than $\pi/6$ or equivalently periods greater than 12 months–one should use at least a 6-month moving average of the 12-month growth rate, but no more than a 12-month moving average. The average value of the filter function using the 6-month moving average over this frequency range is about 1.5, whereas the average value using the 12-month moving average is 0.8; and over higher, less interesting frequencies, the average values are near zero.[13] And with quarterly data, the examples in the top panel of chart 3 suggest that average distortion is minimized using a 5-quarter moving average of the 4-quarter growth rate; this is because the average value of the filter function for frequencies less than $\pi/3$ is 1.25 using a 4-quarter moving average and is 0.6 using a 6-quarter moving average. These computations are consistent with the analytic solution values in the proof to proposition 10, where it is shown that the average value for a 5-quarter moving average is 0.87.

We now turn our attention to the timing properties of combined filters. As noted above, trailing moving averages tend to lag the underlying data (proposition 8), and growth rates tend to lead the underlying data (proposition 5). This implies that at least some of the lagging introduced by a moving average to an underlying series will be undone by a growth rate filter, and vice versa. Indeed, the leads and lags can largely cancel when combining the separate filters; as a result, the combined filter can produce a time series that is in close synchronization with the underlying data and, in so doing, help preserve the underlying timing relationship between completely different time series (for example, capital goods orders and shipments) that separately have been transformed by the same combined filter. The combined filter does not have to produce close synchronization between a series and its filtered value, but quite often does in practice. The phase of the combined filter function is presented in the following proposition.

---

13. For purposes of computation in figures 2 and 3, the horizontal axis—measured from 0 to π radians per unit time—in each case is divided into intervals of size $\pi/100$.

Proposition 11:   The phase of the n-term trailing moving average (with uniform weights) of the k-period growth rate is given by (with summations over j from 0 to n-1 and $\omega \neq 0$):

$$\tan^{-1}\{[-\textstyle\sum_j \sin\omega j + \sum_j \sin\omega(k+j)]/[\sum_j \cos\omega j - \sum_j \cos\omega(k+j)]\}.$$

Proof: Let $x_t = (1/n)[(y_t - y_{t-k}) + (y_{t-1} - y_{t-k-1}) + ... + (y_{t-n+1} - y_{t-k-n+1})]$

$$= (1/n)\{\textstyle\sum y_{t-j} - \sum y_{t-k-j}\}$$

where the summations are over j from 0 to n-1. The frequency response function is:

$$H(\omega) = (1/n)[e^{-i\omega 0} + e^{-i\omega} + ... + e^{-i\omega(n-1)} - e^{-i\omega k} - e^{-i\omega(k+1)} - ... - e^{-i\omega(k+n-1)}]$$

$$= (1/n)\{[1 + \cos\omega + ... + \cos(n-1)\omega - \cos\omega k - \cos\omega(k+1) - ... - \cos\omega(k+n-1)]$$

$$-i[\sin\omega + ... + \sin\omega(n-1) - \sin\omega k - ... - \sin\omega(k+n-1)]\}.$$

Denote the real part of $H(\omega)$ by $R(\omega)$ and the imaginary part by $iI(\omega)$; then, by definition, the phase is $\tan^{-1}[I(\omega)/R(\omega)]$, and the proof is completed by substitution.

As a practical example, the bottom panel of chart 3 compares various filtered values of real GDP in the United States, using quarterly National Income and Product Account (NIPA) data for the period from the first quarter of 1968 to the third quarter of 2001. The shaded areas in the chart denote business cycle peak-to-trough periods as officially designated by the NBER (the cyclical peak in the first quarter of 2001 is not designated on the chart); the NBER peak and trough quarters are quite similar to the corresponding peak and trough quarters of real GDP, differing at most by a quarter. The level of real GDP is a typical macroeconomic time series that is not stationary, based on the augmented Dickey-Fuller test;[14] however, the various filtered series shown are stationary. Although the prior theoretical results are based on an underlying stationary time series process, the results will be seen to hold for real GDP, raising the possibility that they will also hold for other typical macroeconomic time series.

The bottom panel illustrates that the 4-quarter growth rate is a smoother series than

---

14. The t-statistic for the coefficient of the lagged level of real GDP in the augmented Dickey-Fuller regression is 1.4, with critical value of about 2.9 at the 5 percent level.

the annualized 1-quarter growth rate, as suggested by proposition 2 (indeed, the sample variance is reduced from 17.8 to 7.5). Moreover, consistent with proposition 5, the 4-quarter growth rate either moves in synch with the annualized 1-quarter growth rate or lags it; the latter is illustrated during the recession of 1990 when the steep plunge and subsequent bounceback in the 1-quarter growth rate precede that of the 4-quarter growth rate. Over the sample period shown, the annualized 1-quarter growth rate leads the 4-quarter growth rate by about 1 quarter on average prior to cyclical turning points.[15]

The bottom panel also illustrates that the 4-quarter moving average of the 4-quarter growth rate (with a sample variance of 5.5) is a smoother series than the 4-quarter growth rate (proposition 6) and also tends to lag it (proposition 8), the latter being most evident around business-cycle turning points. Further, as suggested by proposition 11, the 4-quarter moving average of the 4-quarter growth rate should be in better synchronization with movement in the underlying series than the 4-quarter growth rate. Indeed, this better synchronization is evident around turning points. For example, the 4-quarter moving average of the 4-quarter growth rate turns down closer in time to the business cycle peaks in 1969, 1973, 1980, 1990, and 2001 than the 4-quarter growth rate but a bit further away in time to the peak in 1981. Interestingly, though, the 4-quarter moving average of the 4-quarter growth rate turns up a bit further away in time to troughs in GDP than the 4-quarter growth rate. However, on average around all business-cycle turning points, the 4-quarter moving average of the 4-quarter growth rate is in better time synchronization with real GDP than is the 4-quarter growth rate. Finally, as established by proposition 10, the combined 4-quarter moving average of the 4-quarter growth rate is better than either of its components at bringing out features of the data associated with business-cycle and lower frequencies, because the average value of the associated filter function is closer to one.

Additional examples of these points are contained in chart 4, which compares various filtered values of the core personal consumption expenditures (PCE) price index using monthly NIPA data for the period from January, 1968 to September, 2001. In

---

15. Although not shown, and consistent with proposition 8, the 4-quarter moving average of real GDP lags GDP at both peaks and troughs generally by 1 to 2 quarters.

particular, it is seen that the 12-month percent change (with variance of 5.0) is a much smoother series than the annualized 3-month percent change (with variance of 5.6), while the 6-month moving average of the 12-month change is smoother still (with variance of 4.9); moreover, the lead introduced by the 12-month growth rate (not shown) is offset by the lag associated with the moving average.
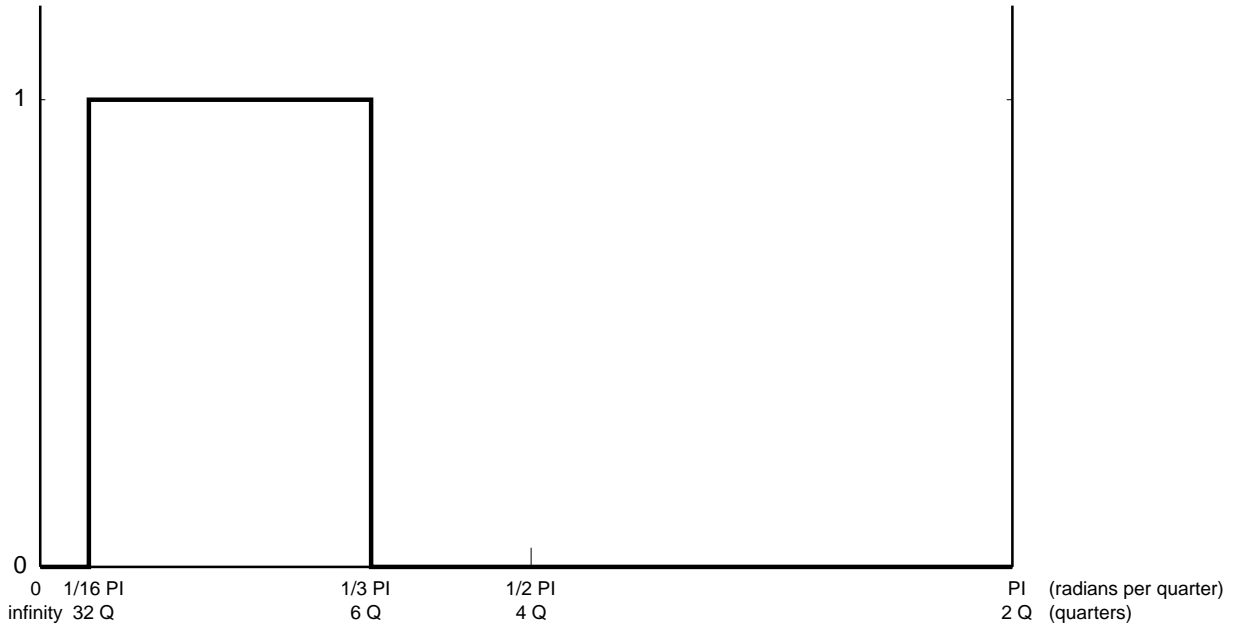
VIII. Conclusion

This paper has examined properties of standard data transformations—in particular, growth rates and moving averages—commonly used in applied economics work. In contrast to recently-developed approximate ideal filters that drop the most recent few years worth of observations from the sample, the filters analyzed here do not drop observations, reflecting common practice among government and financial-market economists. Although these filters are not "ideal," our analysis suggests that the biases and timing shifts introduced by standard data transformations can be substantially neutralized by relying on moving averages of multi-period growth rates.

20

References

Christiano, Lawrence, and Terry Fitzgerald (1999). "The Band Pass Filter," National Bureau of Economic Research Working Paper No. 7257.

Baxter, Marianne, and Robert King (1999). "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, 81, November, pp. 575-593.

Chatfield, Chris (1996). The Analysis of Time Series, 5th edition, Chapman & Hall, London.

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

_____ and Paul Newbold (1986). Forecasting Economic Time Series, 2nd edition. Academic Press, New York.

Hamilton, James (1994). Time Series Analysis. Princeton University Press, Princeton, New Jersey.

Priestley, M.B. (1981). Spectral Analysis and Time Series. Academic Press, New York.

Chart 1

Ideal Business Cycle Filters

Quarterly Data

1				
0				

0   1/16 PI   1/3 PI   1/2 PI   PI   (radians per quarter)
infinity 32 Q   6 Q   4 Q   2 Q (quarters)

Monthly Data

1

0

1/48 PI   1/9 PI 1/6 PI   PI   (radians per month)
96 M   18 M   12 M   2 M (months)
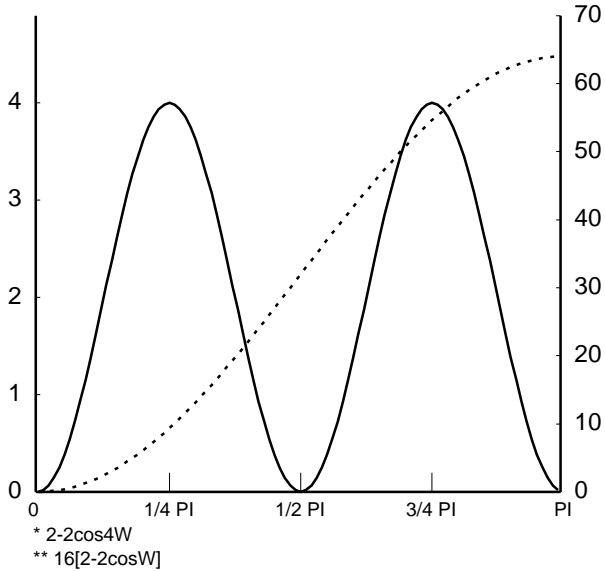
Chart 2

Filter Functions of Data Transformations

Percent Changes Typically Used With Quarterly Data
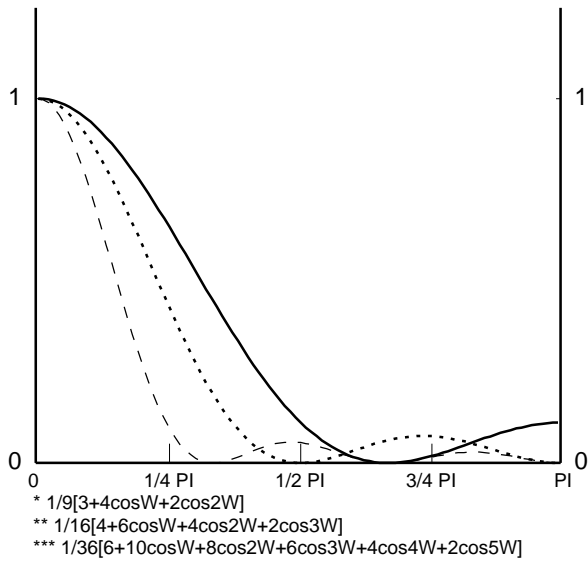
— 4-quarter* (left)
···· 1-quarter, a.r** (right)

0   1/4 PI   1/2 PI   3/4 PI   PI

* 2-2cos4W
** 16[2-2cosW]

Percent Changes Typically Used With Monthly Data

— 12-month* (left)
···· 3-month, a.r** (right)

0   1/6 PI   1/3 PI   1/2 PI   2/3 PI   5/6 PI   PI

* 2-2cos12W
** 16[2-2cos3W]
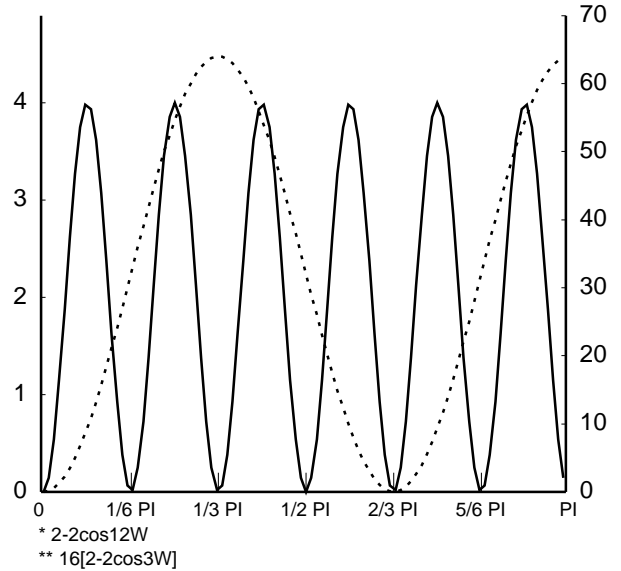
Moving Averages

— 3 period*
···· 4 period**
– – 6 period***

0   1/4 PI   1/2 PI   3/4 PI   PI

* 1/9[3+4cosW+2cos2W]
** 1/16[4+6cosW+4cos2W+2cos3W]
*** 1/36[6+10cosW+8cos2W+6cos3W+4cos4W+2cos5W]

Combined Moving Average and Twelve-Month Growth Rate

— 3 period*
···· 4 period**
– – 6 period***

0   1/6 PI   1/3 PI   1/2 PI   2/3 PI   5/6 PI   PI

* [2-2cos12W]1/9[3+4cosW+2cos2W]
** [2-2cos12W]1/16[4+6cosW+4cos2W+2cos3W]
*** [2-2cos12W]1/36[6+10cosW+8cos2W+6cos3W+4cos4W+2cos5W]
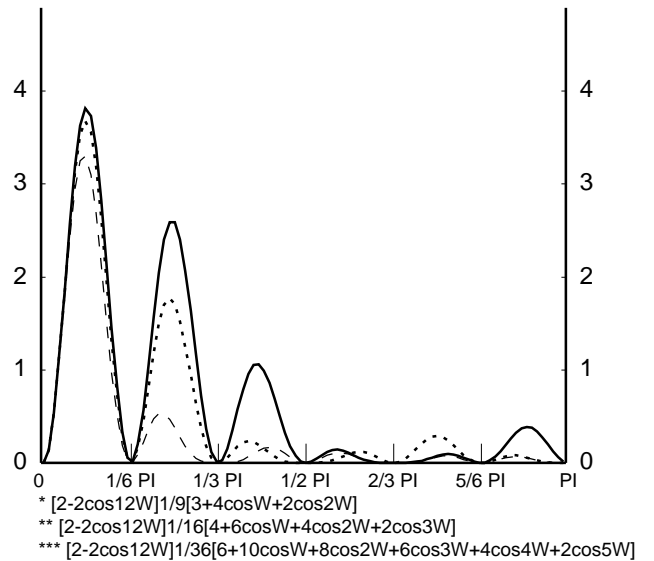
Chart 3

Combined Moving Average and Four-Quarter Growth Rate

————— 3 period*

············· 4 period**

— – — – 6 period***

0   1/4 PI   1/2 PI   3/4 PI   PI

 * [2-2cos4W] 1/9[3+4cosW+2cos2W]

 ** [2-2cos4W] 1/16[4+6cosW+4cos2W+2cos3W]

 *** [2-2cos4W]1/36[6+10cosW+8cos2W+6cos3W+4cos4W+2cos5W]

Real GDP
(Percent change)

————— 4-quarter

············· 1-quarter, a.r.

————— 4-quarter moving average of 4-quarter
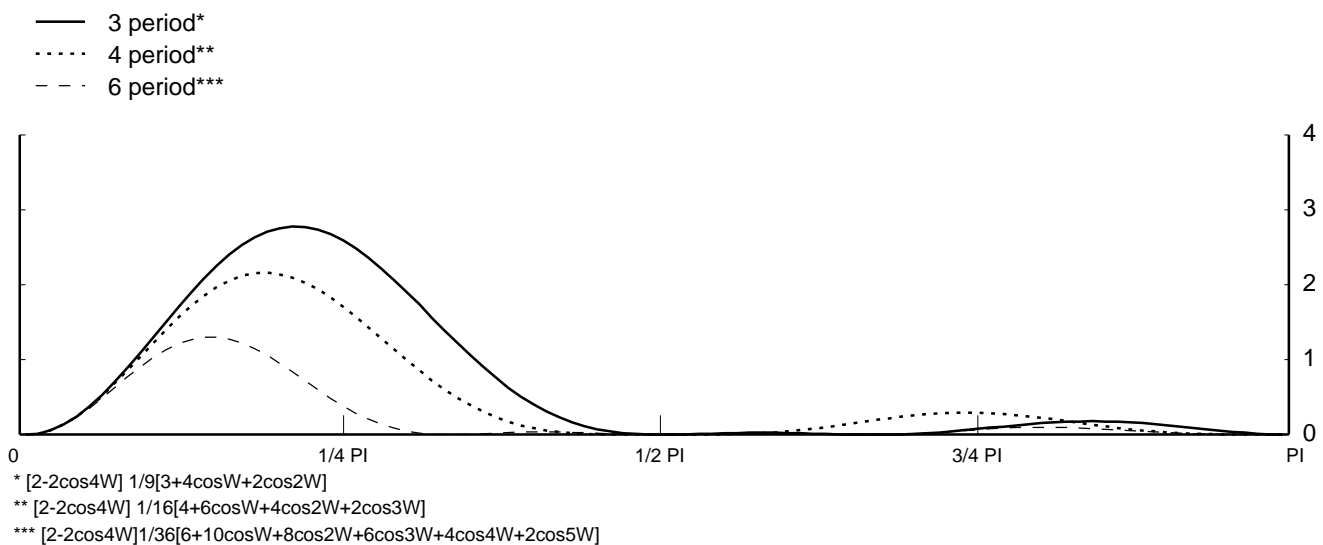
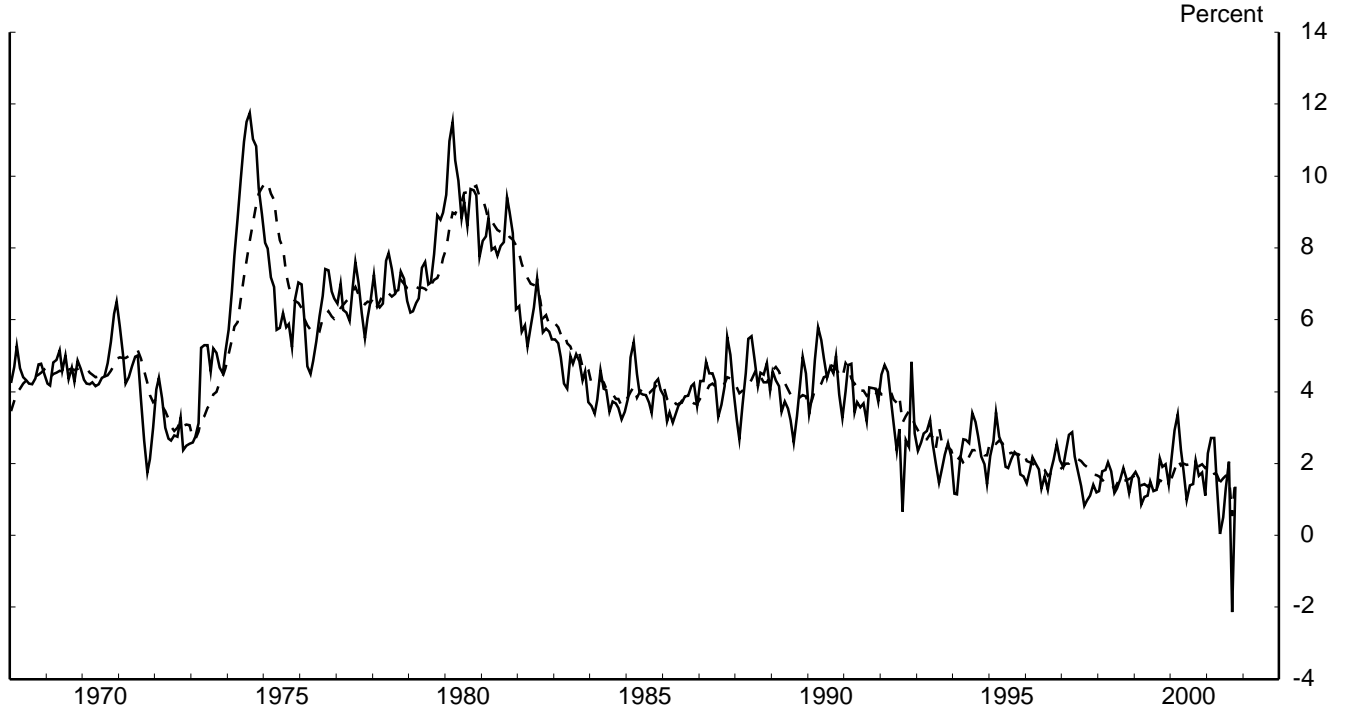————— Level, billions of dollars, left scale
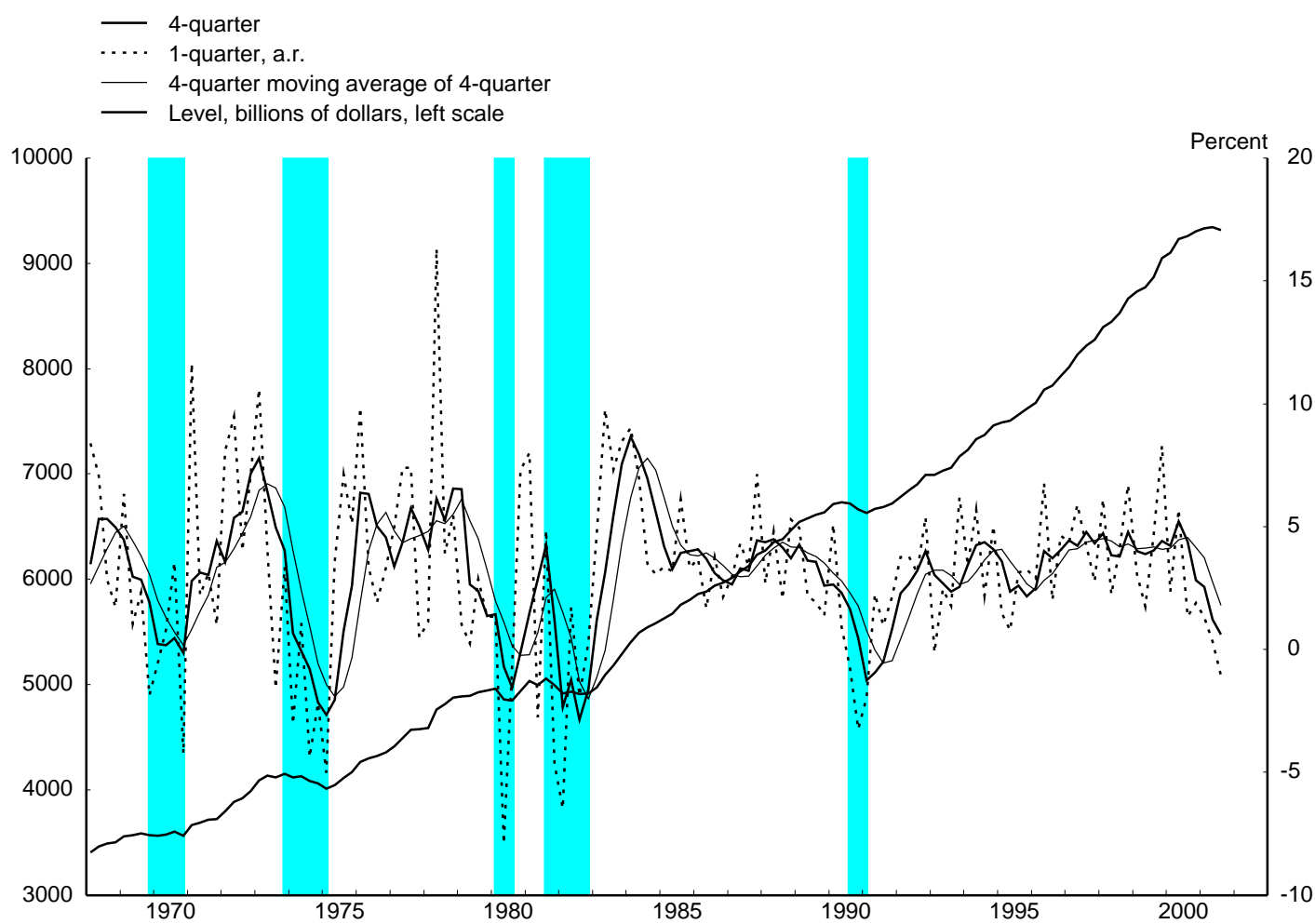
Note. Shaded regions represent NBER business cycle peak-to-trough periods.

Chart 4

Core PCE Prices
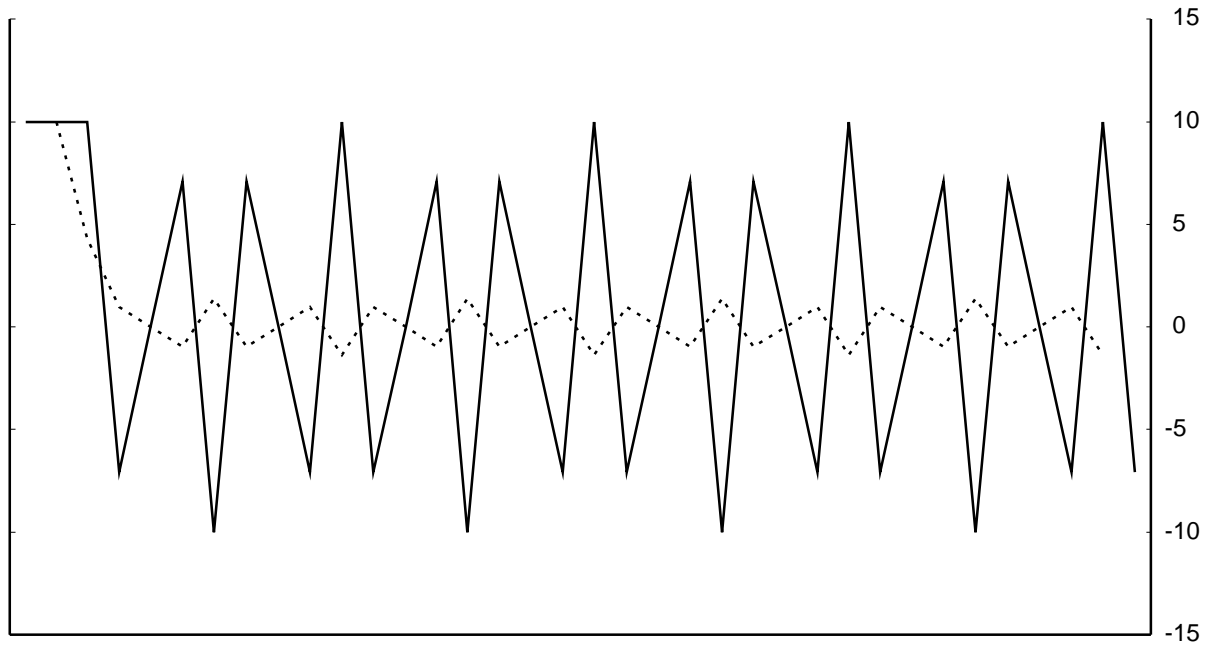**(Percent change)**

—— 3-month, a.r.
– – · 12-month

Percent

14

12

10

8

6

4

2

0

-2

-4

1970   1975   1980   1985   1990   1995   2000

—— 12-month
– – · 6-month moving average of 12-month

Percent

12

10

8

6

4

2

0

1970   1975   1980   1985   1990   1995   2000
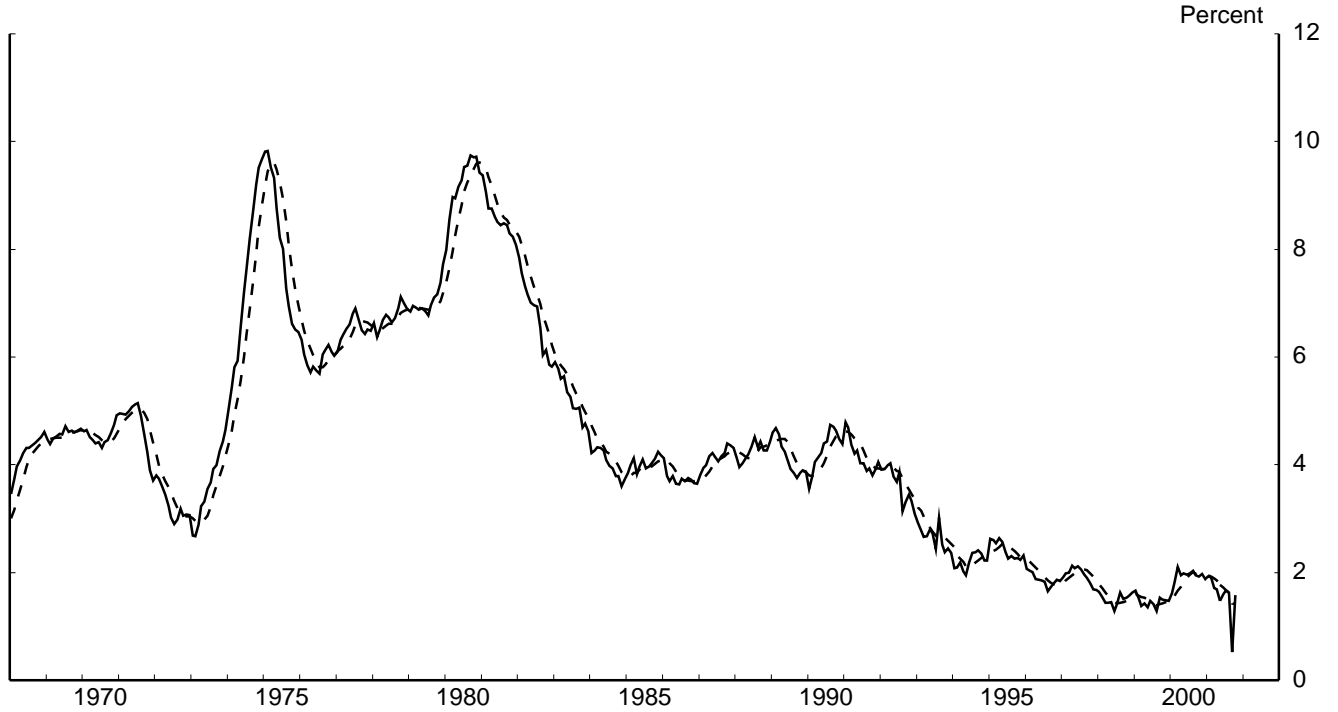
Appendix 1

3-term centered moving average

—— 10cos[3/4 PI t]
· · · · 3-term centered moving average

4-term trailing moving average

—— 10cos[3/4 PI t]
· · · · 4-term trailing moving average