Hitotsubashi Journal of Commerce and Management 35 (2000) pp.19-47.

ON FINANCIAL TIME SERIES DECOMPOSITIONS WITH APPLICATIONS TO VOLATILITY

KJELL DOKSUM*, RYOZO MIURA**, AND HIROAKI YAMAUCHI***

Abstract

We consider decompositions of financial time series that identify important modes of variation in the series. The first term in the decomposition measures long-term trends and focuses on large-scale features of variability. The second term measures short-term trends and local features of variability remaining after the long-term trend has been removed. The third term measures the irregularity left in the series after the long- and short-term trends have been subtracted out. This term is further broken down by regressing it on its own lagged values. One goal of this decomposition is to transform a "raw" time series into three interpretable terms plus a term that is approximately noise. In this paper, the methodology is applied to the exchange rates of Japanese Yen (JY), German Marks (GM), Swiss Francs (SF), and British Pounds (BP) in the unit of U.S. dollars. Similarities and differences in the trends between these currencies as well as their volatilities are discussed.

Keywords: Autoregressive fit, locally linear regression, decomposition, exchange rates, financial time series, LOWESS, volatility
JEL classification: C10, C14

I. *Introduction*

In this paper we decompose financial time series by using and extending the ideas of Shibata and Miura (1997) who introduced a two-step smoothing technique. We consider the properties of such decompositions and apply them to currency exchange rate time series.

The initial decomposition of the time series $Y(t)$, $t=1, \ldots, T$, is of the from

$$Y(t)=L(t)+S(t)+I(t)$$

where $L(t)$ represents "long-term trend", $S(t)$ stands for "short-term trend", and $I(t)=Y(t)-[L(t)+S(t)]$ is the "residual" or "irregular" part of the series. More precisely, we assume that $Y(t)$ is a random variable with population mean $\mu(t)$, $t=1, \ldots, T$, and define $L(t)$, the

* Department of Statistics, University of California, Berkeley; doksum@stat.berkeley.edu Partly supported by NSF grant DMS-9971301.
** Graduate School of International Corporate Strategy, Hitotsubashi University; rmiura@ics.hit-u.ac.jp
*** MTB Investment Technology Institute Co.,Ltd; yamauchi@mtec-institute.co.jp

theoretical long-term trend, as a weighted average of $\mu(t)$ over a long time interval such as a quarter of a year. After subtracting out the long-term trend, the short-term trend $S(t)$ is defined as a short-term weighted average of $\mu(t)-L(t)$. The residual can now be measured by the series $I(t)=Y(t)-[L(t)+S(t)]$, which we call the theoretical residual or irregular series.

The empirical decomposition $Y(t)$ is defined as

$$Y(t)=\hat{L}(t)+\hat{S}(t)+\hat{I}(t)$$

where $\hat{L}(t)$ and $\hat{S}(t)$ are the empirical weighted average obtained by replacing $\mu(t)$ by $Y(t)$ in the definitions of $L(t)$ and $S(t)$, and $\hat{I}(t)=Y(t)-[\hat{L}(t)+\hat{S}(t)]$.

Our decomposition is similar to the analysis of variance decomposition, which decompose a variance into a sum of variability due to signal plus variability due to noise except we have decomposed the signal into two parts, large time-scale and short time-scale variability. Moreover, because of the dependence over time in the series, we will further decompose the noise term $I(t)$ when necessary into terms related to lagged values plus "pure" noise using parametric autoregressive techniques.

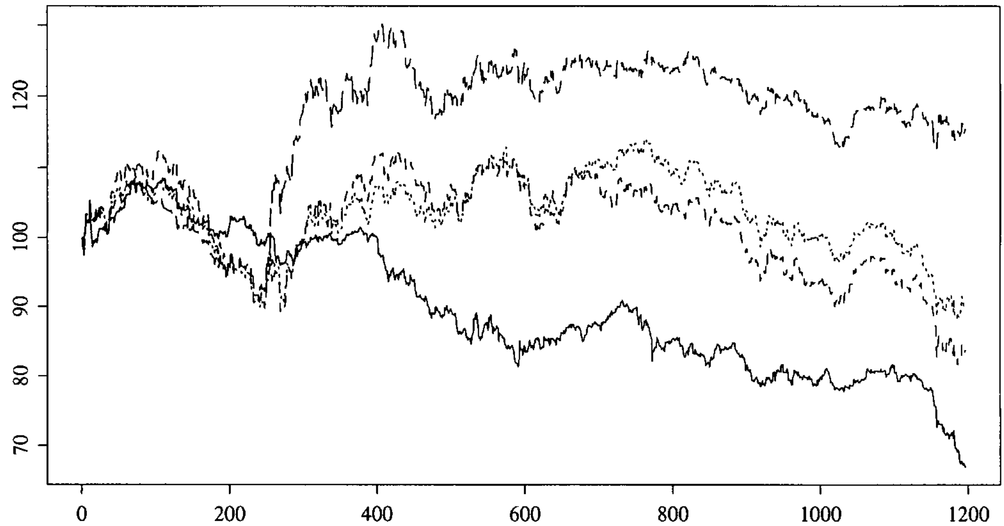
Our approach uses nonparametric regression techniques, on one hand, to model long- and short-term trends; and it uses classical parametric time-series techniques, on the other hand, to model the residual $I(t)$ of the fit $L(t)+S(t)$.

We apply the decompositions to the analysis of exchange rate data; in particular, to the exchange rates of a U.S. dollar in units of JY, GM, SF, and BP for the 1196 days from January 5, 1992, until April 14, 1995. The exchange rates are converted to the percentage value of the initial value (January 5, 1992) before applying the decomposition (See Figure 1.1.). We find that the GM and SF series are very similar in that not only do they have close long-term trends, but their short-term trends and residual series are also very close. We use our decomposition to give graphical representation of the relationship between GM and SF exchange rates. We conjecture that the similarity between the movements of GM and SF is related to the macro-economic relations with the U.S. economy and also related to the attitude of currency traders who simultaneously trade the SF in the same direction as the GM. There could be a perception by dealers that when one of these two currencies moves, the other will move in the same direction; so if one of them falls behind, it will quickly catch up to the other one as dealers try to take advantage of the gap.

Starting about 600 days into the time period, the long-term trend of JY becomes very similar to that of GM and SF. After 720 days into the period, the dollar shows a steady long-term decline against JY, GM, and SF; but the decline is not linear. It switches between concave and convex. This represents a macro-level economic relationship to the U.S. dollar.

We consider two types of decompositions. The first type is a decomposition as in Shibata and Miura (1997) and is based on an ex-post analysis of the data in the time series. Its decomposition is symmetric in the sense that it gives equal weight to values before and after each time point t. It gives a description based on all of the global and local tendencies of the time series. The second type of smoothing, which is often used for the prediction of a time series, is called here *predictive* since it is based on values of the time series up to the time t. In Sections II and III we explain the details of the decomposition methodologies with symmetric and predictive smoothing respectively and display the results of its applications to the four currency exchange rates. In Section IV we discuss the implications of our decomposition results for risk measurements and derivative pricing based on VaR (Value at Risk).

FIGURE 1.1.
Indexed Exchange Rates: JY, GM, SF, and BP



Indexed currency exchange rates of the JY ———, GM ·········, SF - - - -, and BP – – – for the time period January 5, 1992, to April 14,1995.

In Section V we give an application of the predictive decomposition to the assessment of weekday volatility in the exchange rates. Our approach is based on introducing $V(t)=[X(t)-m(t)]^2$, where $X(t)=\log[Y(t)/Y(t-1)]$ are returns and $m(t)$ is the conditional mean of $X(t)$ given the previous values $X(t-r)$, $r\geq 1$. An estimate of heterogeneous volatility based on previous values $V(t-r)$, $r\geq 1$, of $V(t)$ can then be obtained from the estimated sum $\hat{L}_{PV}(t)+\hat{S}_{PV}(t)$ of the predictive long- and short-term trends of $V(t)$. We find that the volatilities of GM and SF are close, while the JY follows a more independent path. For the Japanese series we compare our method to J. P. Morgan's RiskMetrics™ (1994).

In Section VI we derive asymptotic approximations to the long- and short-term trends and show how they are affected by sudden changes in the mean of the time series. We also show that the locally linear predictive methods introduced in Section III tend to represent the "true" trend better than kernel methods such as J. P. Morgan's RiskMetrics™ whenever the methods are applied to series where trends are present.

This paper is a shortened version of Doksum, Miura, and Yamauchi (1998).

II. *Symmetric Decomposition of Time Series*

(i) Long-Term Trend

We define $L(t)$ as a weighted average of means over a long time span. That is

$$L(t)=\sum_{k=t-M}^{t+M} w_k(t)\,\mu(k) \tag{2.1}$$

for appropriate weights that nearly sum to one; examples of such weights are the kernel weights

$$w_k(t) = \frac{1}{M} K\left(\frac{k-t}{M}\right), \text{ for } t-M \leq k \leq t+M \tag{2.2}$$

where the kernel $K(u)$ is a symmetric density on the interval $[-1, 1]$.

A very simple and easily understood set of weights are the "uniform moving average weights", which equal $(2M)^{-1}$ for k in the interval $[t-M, t+M]$ and equal 0 elsewhere. In this case $L(t)$ is the uniform moving average $L(t) = (2M)^{-1} \sum_{k=t-M}^{t+M} \mu(t)$.

Another interesting set of weights $w_k(t)$ is given by the set of "exponential weights" $w_k(t) = \dfrac{\lambda^{|t-k|}}{\sum_{k=-M}^{M} \lambda^{|t-k|}}$, $0 < \lambda < 1$. These weights are symmetric versions, standardized to sum to one, of the weights used by J. P. Morgan's RiskMetrics^{TM} (1994), who recommends $\lambda = 0.94$. The related weights $(1-\lambda)\lambda^{(t-k)}$ are from the Integrated Moving Average process ARIMA (0, 1, 1). See Box and Jenkins (1976), page 105.

We also consider locally linear or locally polynomial weights $w_k(t)$. These are defined as follows: Let $w_k^*(t)$ be a preliminary set of nonnegative weights such as (2.2), and let $a(t)$, $b_1(t)$, ..., $b_d(t)$ be the values of α, β_1, ..., β_d that minimize the local weighted sum of squares

$$\sum_{k=t-M}^{t+M} \left[\mu(k) - \left\{ \alpha + \sum_{j=1}^{d} \beta_j (k-t)^j \right\} \right]^2 w_k^*(t). \tag{2.3}$$

Then $\tilde{\mu}(s) = a(t) + \sum_{j=1}^{d} b_j(t)(s-t)^j$ is called the locally polynomial fit, and the locally polynomial long-term trend at time t is $L(t) = \tilde{\mu}(t) = a(t)$. It can be shown that $L(t)$ can be written in the form

$$L(t) = a(t) = \sum_{k=t-M}^{t+M} w_k(t) \mu(k). \tag{2.4}$$

where $w_k(t)$ are constants; e.g., Fan and Gijbels (1996), page 20.

Locally polynomial smoothers of $\mu(t)$ have many desirable properties that have been demonstrated by Cleveland (1979), Cleveland and Devlin (1988), Fan (1993), Fan and Gijbels (1996), and Ruppert and Wand (1994), among others. The regression smoother LOWESS in S-PLUS (1993) is a locally linear smoother adjusted for outliers (extreme values) among the $\{\mu(t)\}$. LOESS, which also is in S-PLUS, includes locally polynomial smoothers. The locally polynomial long-term trend function $L(t)$ has the important *polynomial reproduction property*. That is, if the true mean $\mu(t)$ is a polynomial of degree p, then the degree $d=p$ locally polynomial long-term trend function $L(t)$ equals $\mu(t)$ for all $t=1, 2, ..., T$. See, e.g., Fan and Gijbels (1996), page 61.

The set of weights should be chosen to balance between stability, familiarity, and interpretability. For instance, a 91-day uniform moving average $L(t)$ can be interpreted as focusing on quarterly trends. Moreover, such moving averages of exchange rates are of particular interest because they are used in place of daily rates in many financial instruments. On the other hand, moving averages are unstable in periods of sudden changes at the beginning and end of a series (the boundary curse). A much more stable measure of long-term trend is obtained by using the locally linear $L(t)$ defined by (2.3) and (2.4) with $d=1$.

The long-term trend can be used to analyze the development of a series without the distractions of local time changes, and it is especially useful for comparing similar time series. To illustrate, we introduce the natural unbiased estimator of $L(t)$, which is

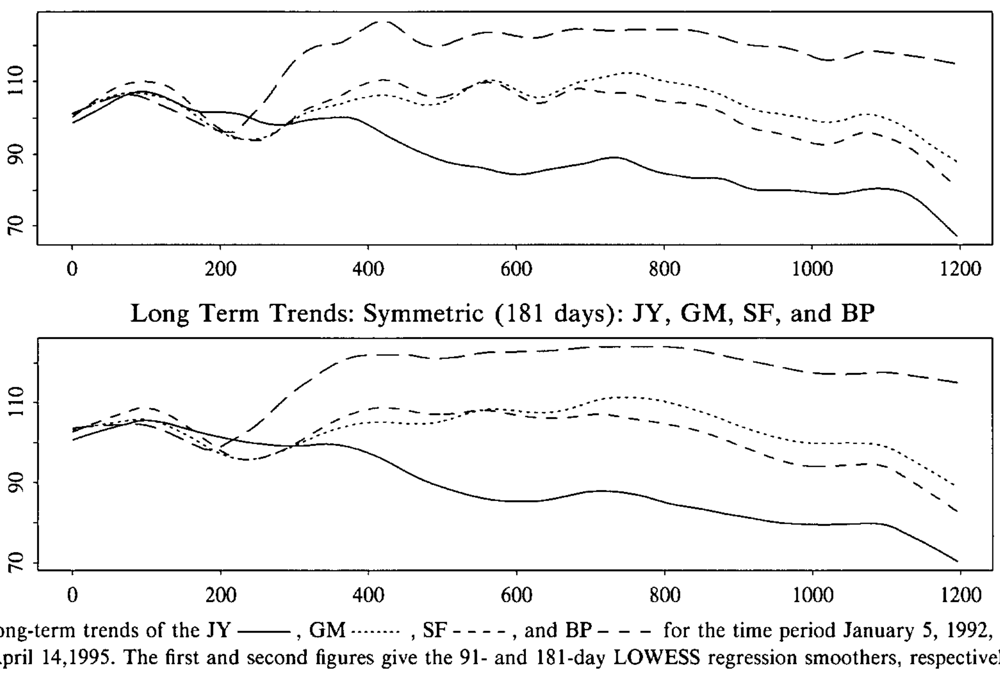
$$\hat{L}(t)=\sum_{k=t-TL}^{t+TL} w_k(t)\,Y(k).$$

Figure 2.1 gives the estimated long-term trends as represented by the 91- and 181-day locally linear LOWESS smoothers for the exchange rates of JY, GM, SF, and BP in U.S. dollars for the time period January 5, 1992, until April 14, 1995. There are 1196 days (including weekends) in this time period during which there were, respectively, 36, 34, 51, and 40 missing values (holidays) among the JY, GM, SF, and BP. The missing values are replaced by the linear interpolate between the nearest two actual values in the series. This has very little effect on the regression decomposition since $\hat{L}(t)$ is a "smoother", which smoothes out the effect of individual points. The four exchange rates series all have been rescaled to have initial value 100. That is, the graphs show the percentage value of a U.S. dollar in the four currencies when compared to the value on January 5, 1992. The four exchange rate values were all recorded at the same instance, 10 p.m., GMT.

The graphs show that SF and GM have very similar long-term trends while the long-term trends in JY and BP follow their own low and high paths, respectively. The "spike" in the BP shortly after day 200 corresponds to the time when it was "cut loose" from other European

FIGURE 2.1.

Long Term Trends: Symmetric (91 days): JY, GM, SF, and BP



Long-term trends of the JY ——, GM ········ , SF - - - - , and BP – – – for the time period January 5, 1992, to April 14,1995. The first and second figures give the 91- and 181-day LOWESS regression smoothers, respectively.

currencies. Also note that starting near day 600, the path of the JY is very similar to that of the GM and SF. The 91- and 181-day long-term trends are very similar except for the BP shortly after day 200 when it was "cut loose" from other European currencies.

The correlations between the long-term trends of the four currencies are given in Table 2.1. It gives a concise summary of the trends shown in the graphs. It also provides one surprise: The correlations of the long-term trend between JY and SF is much higher than that between JY and GM, indicating a surprisingly strong macro-economic relationship between the Japanese and Swiss currencies.

TABLE 2.1.

(a) 91 days

	JY	GM	SF	BP
JY	1.00	0.11	0.52	−0.59
GM	0.11	1.00	0.85	0.52
SF	0.52	0.85	1.00	0.24
BP	−0.59	0.52	0.24	1.00

(b) 181 days

JY	GM	SF	BP
1.00	0.09	0.53	−0.66
0.09	1.00	0.85	0.51
0.53	0.85	1.00	0.18
−0.66	0.51	0.18	1.00

Correlations between long-term trends in the four currencies. Tables (a) and (b) give, respectively, the correlations for LOWESS 91-day and 181-day long-term trends.

(ii) Short-Term Trend

We define the short-term trend $S(t)$ as a weighted average over a short time span of deviations of the long-term trend from the mean. That is

$$S(t)=\sum_{k=t-TS}^{t+TS} \nu_k(t)[\mu(k)-L(k)]$$

where the weights $\nu_k(t)$ are nonzero over a short time span. The trend function $S(t)$ focuses on changes over short time periods, such as one or two weeks, and gives short term "micro" fluctuations that are not measured by the long term "macro" trend function $L(t)$. In the case of exchange rates where there are strong "day of the week" effects (currencies are traded seven days a week, but Saturday and Sunday trades are subject to conditions different from the other days of the week), the seven-day uniform moving average is without the day-of-the-week effect.

The natural unbiased estimator of $S(t)$ is

$$\hat{S}(t)=\sum_{k=t-TS}^{t+TS} \nu_k(t)[Y(k)-\hat{L}(t)].$$

Figure 2.2 gives the estimated short-term trends $\hat{S}(t)$ based on locally linear LOWESS smoothers for the four currencies we consider. We see that the GM and SF have very similar short-term trends, while the JY and BP do not. The correlations are given in Table 2.2.

The relatively small correlations between JY and the other currencies is an indication that its short-term movements are more influenced by national economic developments than the other currencies. Moreover, a strong local relationship between SF and GM is evident in the correlation tables.

TABLE 2.2.

(a) (91,7) days

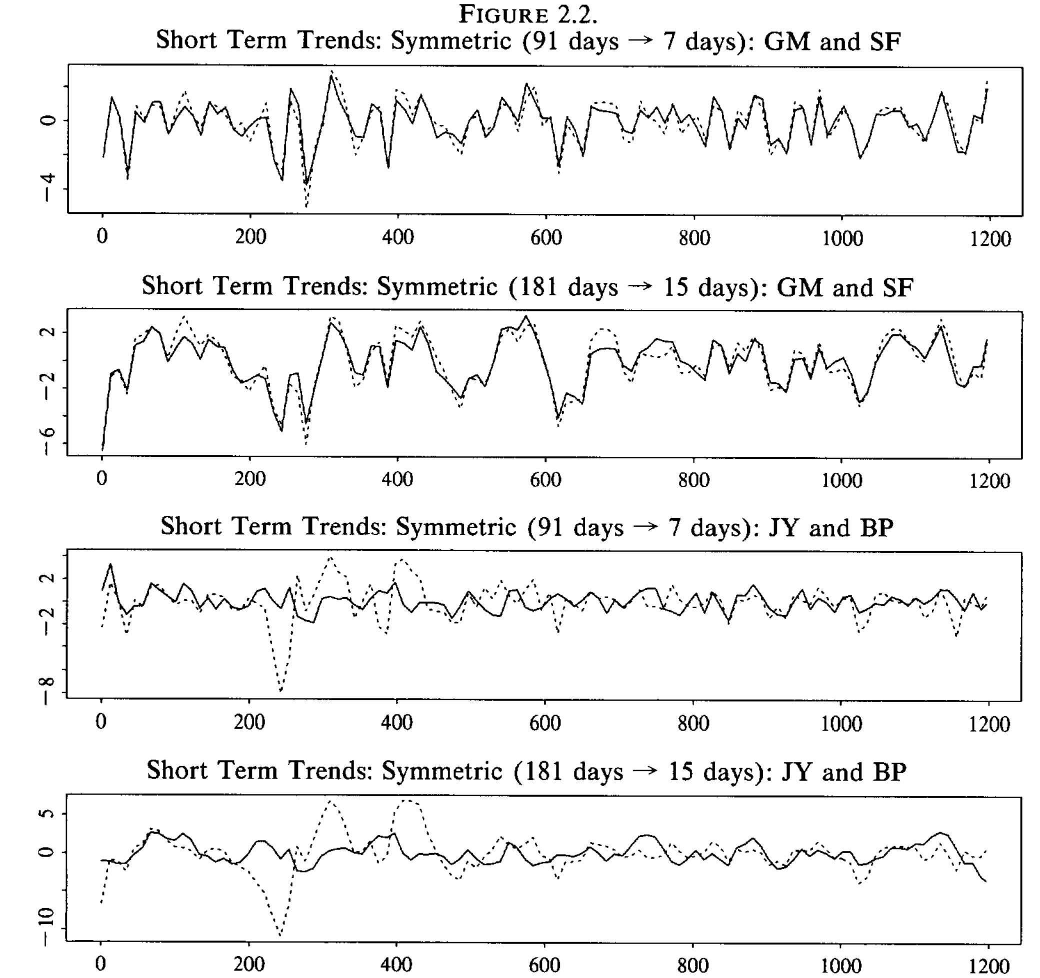
	JY	GM	SF	BP
JY	1.00	0.41	0.43	0.10
GM	0.41	1.00	0.93	0.64
SF	0.43	0.93	1.00	0.63
BP	0.10	0.64	0.63	1.00

(b) (181,15) days

JY	GM	SF	BP
1.00	0.42	0.46	0.15
0.42	1.00	0.95	0.71
0.46	0.95	1.00	0.70
0.15	0.71	0.70	1.00

Correlations between short-term trends in the four currencies. (a) uses a 7-day short-term trend based on a 91-day long-term trend while (b) uses a 15-day short-term trend based on a 181-day long-term trend.

FIGURE 2.2.



The first two graphs give the short-term trends of GM —— and SF ------ using, respectively, 7- and 15-day short-term trends based, respectively, on 91- and 181-day long-term trends. The next two graphs similarly give short-term trends of JY —— and BP ------ .

(iii) The Residual Series

We define the residual or irregular series $I(t)$ as the residual of the original series $Y(t)$ after subtracting the sum $L(t)+S(t)$ of the long- and short-term trends. That is

$$I(t)=Y(t)-[L(t)+S(t)].$$

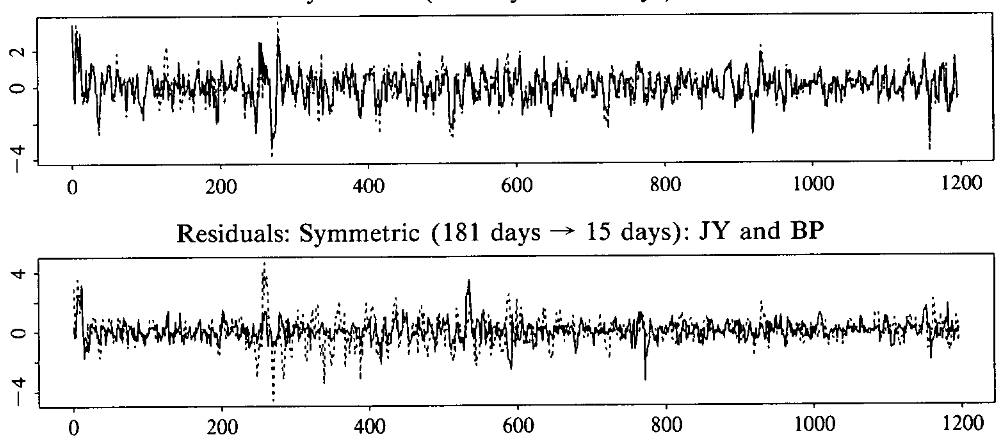
Our empirical residual series is

$$\hat{I}(t)=Y(t)-[\hat{L}(t)+\hat{S}(t)].$$

Figure 2.3 gives the empirical residual series of the four currencies when the long- and short-term trends are based on 181- and 15-day spans, respectively. The (91,7) day empirical residuals were very similar (not shown here). The same story as in Section II(i) and II(ii) emerges: The GM and SF are closely correlated while the JY and BP show less of a relationship to the other currencies. The BP still shows a spike after it was cut loose from other European currencies. The correlations are given in Table 2.3.

FIGURE 2.3.

Residuals: Symmetric (181 days → 15 days): GM and SF



The first graph gives the residual series for GM ——— and SF --------- using 181- and 15-day long- and short-term trends. The next graph similarly gives the residual series for the JY ——— and BP --------- .

TABLE 2.3.

(a) (91,7) days

	JY	GM	SF	BP
JY	1.00	0.47	0.46	0.33
GM	0.47	1.00	0.91	0.72
SF	0.46	0.91	1.00	0.68
BP	0.33	0.72	0.68	1.00

(b) (181,15) days

JY	GM	SF	BP
1.00	0.50	0.48	0.38
0.50	1.00	0.92	0.75
0.48	0.92	1.00	0.69
0.38	0.75	0.69	1.00

Correlations between residuals of the four currencies. (a) uses a 7-day short-term trend based on a 91-day long-term trend while (b) uses a 15-day short-term trend based on a 181-day long-term trend.

(iv) The Decompositions

Our theoretical and empirical decompositions of the time series $Y(t)$ are

$$Y(t)=L(t)+S(t)+I(t), \text{ and } Y(t)=\hat{L}(t)+\hat{S}(t)+\hat{I}(t).$$

An interesting question is whether the three series $\hat{L}(t)$, $\hat{S}(t)$, and $\hat{I}(t)$ are nearly orthogonal or uncorrelated. The correlation between them measures the extent to which each series provides "orthogonal information". For instance, we hope that the short-term trend $\hat{S}(t)$ is nearly orthogonal to the long-term trend $\hat{L}(t)$. Moreover, the correlation between $\hat{L}(t)+\hat{S}(t)$ and $\hat{I}(t)$ measure to what extent the decomposition

$$Y(t)-\bar{Y}=[\hat{L}(t)+\hat{S}(t)-\bar{Y}]+\hat{I}(t)$$

have been successful in splitting the variability in the series into a signal and a noise part. The correlations are given in Table 2.4.

The correlations between $\hat{L}(t)$ and $\hat{S}(t)$ in Table 2.4 show that JY is decomposed better than the other currencies, since for JY this correlation is close to zero. For the time spans

TABLE 2.4.

Correlation for (91,7) day trends	JAPANESE YEN				GERMAN MARKS			
	$\hat{L}(t)$	$\hat{S}(t)$	$\hat{L}(t)+\hat{S}(t)$	$\hat{I}(t)$	$\hat{L}(t)$	$\hat{S}(t)$	$\hat{L}(t)+\hat{S}(t)$	$\hat{I}(t)$
$\hat{L}(t)$	1.00	0.08	1.00	0.00	1.00	0.23	0.98	−0.02
$\hat{S}(t)$	0.08	1.00	0.15	−0.04	0.23	1.00	0.39	0.11
$\hat{L}(t)+\hat{S}(t)$	1.00	0.15	1.00	0.00	0.98	0.39	1.00	0.00
$\hat{I}(t)$	0.00	−0.04	0.00	1.00	−0.02	0.11	0.00	1.00
Correlation for (91,7) day trends	**SWISS FRANCS**				**BRITISH POUNDS**			
	$\hat{L}(t)$	$\hat{S}(t)$	$\hat{L}(t)+\hat{S}(t)$	$\hat{I}(t)$	$\hat{L}(t)$	$\hat{S}(t)$	$\hat{L}(t)+\hat{S}(t)$	
$\hat{L}(t)$	1.00	0.19	0.99	−0.03	1.00	0.29	0.99	−0.06
$\hat{S}(t)$	0.19	1.00	0.34	0.07	0.29	1.00	0.43	−0.04
$\hat{L}(t)+\hat{S}(t)$	0.99	0.34	1.00	−0.02	0.99	0.43	1.00	−0.07
$\hat{I}(t)$	−0.03	0.07	−0.02	1.00	−0.06	−0.04	−0.07	1.00

Correlation for (181,15) day trends	JAPANESE YEN				GERMAN MARKS			
	$\hat{L}(t)$	$\hat{S}(t)$	$\hat{L}(t)+\hat{S}(t)$	$\hat{I}(t)$	$\hat{L}(t)$	$\hat{S}(t)$	$\hat{L}(t)+\hat{S}(t)$	$\hat{I}(t)$
$\hat{L}(t)$	1.00	0.09	0.99	0.04	1.00	0.23	0.96	0.00
$\hat{S}(t)$	0.09	1.00	0.22	0.11	0.23	1.00	0.49	0.13
$\hat{L}(t)+\hat{S}(t)$	0.99	0.22	1.00	0.05	0.96	0.49	1.00	0.04
$\hat{I}(t)$	0.04	0.11	0.05	1.00	0.00	0.13	0.04	1.00
Correlation for (181,15) day trends	**SWISS FRANCS**				**BRITISH POUNDS**			
	$\hat{L}(t)$	$\hat{S}(t)$	$\hat{L}(t)+\hat{S}(t)$	$\hat{I}(t)$	$\hat{L}(t)$	$\hat{S}(t)$	$\hat{L}(t)+\hat{S}(t)$	$\hat{I}(t)$
$\hat{L}(t)$	1.00	0.19	0.97	0.01	1.00	0.29	0.96	−0.05
$\hat{S}(t)$	0.19	1.00	0.44	0.13	0.29	1.00	0.53	0.04
$\hat{L}(t)+\hat{S}(t)$	0.97	0.44	1.00	0.05	0.96	0.53	1.00	−0.03
$\hat{I}(t)$	0.01	0.13	0.05	1.00	−0.05	0.04	−0.03	1.00

Correlations between the terms in the decompositions of four exchange rate series.

considered, GM, SF, and BP have a more complicated structure and may require different time spans to achieve orthogonality between $\hat{L}(t)$ and $\hat{S}(t)$. On the other hand, the decompositions have achieved near orthogonality between $\hat{L}(t)+\hat{S}(t)$ and the residuals $\hat{I}(t)$ for all the four currencies.

Finally, we compared (not shown here) the variability of the original time series with that of the terms in the decomposition by giving the box plots of $Y(t)-Y(t-1)$, $\hat{L}(t)-\hat{L}(t-1)$, $\hat{S}(t)-\hat{S}(t-1)$, $\hat{I}(t)-\hat{I}(t-1)$, $\hat{I}(t)$, and the residuals of a vector autoregressive fit to $\hat{I}(t)$. These box plots indicate the frequency distribution of daily movements of the terms in the decomposition. They show that a very large part of the daily increments $Y(t)-Y(t-1)$ of $Y(\cdot)$ come from the irregulars, and the daily increments of $L(t)$ and $S(t)$ are very small. This means that the increment or change of $L(\cdot)$ and $S(\cdot)$ are not so important in the prediction of $Y(\cdot)$, but the change of $I(\cdot)$ is the main action. Thus, the main part of prediction will depend on the (possibly stationary) behavior of $I(t)$.

Remark 2.1. *Comments on the types of decompositions.* Shibata and Miura (1997) discuss differences and similarities between the decomposition $Y(t)=\hat{L}(t)+\hat{S}(t)+\hat{I}(t)$ and the decomposition "SABL" proposed by Cleveland et al. (1981) and implemented by Becker et al. (1988). Other decompositions of time series have been proposed by Beveridge and Nelson (1981). The application of regression smoothing techniques to time series data is treated in monographs by Müller (1988) and Györfi, Härdle, Sarda, and Vieu (1989).

The issue of how long the long time span in $L(t)$ and the short time span in $S(t)$ should be chosen can often be decided by practical considerations. For instance, long time spans of one quarter or one year are natural and short time spans of one or two weeks are reasonable choices. The time spans could also be chosen to minimize the correlations between the terms in the decompositions. Cleveland et al. (1981) proposed using a power transformation of the original series $Y(t)$ to minimize the correlation between the terms in their SABL decomposition. This could also be done with the LAST (long- and short-term) decomposition, in particular, using the nonparametric correlation coefficient of Doksum and Samarov (1995). However, this was not done in our exchange rate analysis since the squared correlations from Table 2.4 are very low (bounded from above by 0.09) and because the original scale (exchange rate) is easier to interpret than a power transformation.

(v) Autoregressive Analysis of Residuals

We investigate the dependence of $\hat{I}(t)$ on its lagged values by using a multivariate autoregressive fit program ("ar" in S-PLUS). We find that the Akaike Information Criteria selects an order 4 autoregressive model. To check the success of the autoregressive fit we applied the BDS test. (See Brock, Dechert and Scheinkman (1987); Brock, Hsieh, and LeBaron (1991).) The results (not shown here) suggest that properties of the residuals are very much like that of an independent and identically distributed sequence of random variables.

The four-variate autoregression model (MAR(4)) is

$$\mathbf{I}(t)=\mathbf{C}_1\cdot\mathbf{I}(t-1)+\mathbf{C}_2\cdot\mathbf{I}(t-2)+\mathbf{C}_3\cdot\mathbf{I}(t-3)+\mathbf{C}_4\cdot\mathbf{I}(t-4)+\mathbf{e}(t), \tag{2.5}$$

where

$$\mathbf{I}(t)=(I^{JY}(t),\ I^{GM}(t),\ I^{SF}(t),\ I^{BP}(t))^T,$$

$$\mathbf{e}(t)=(\epsilon^{JY}(t),\ \epsilon^{GM}(t),\ \epsilon^{SF}(t),\ \epsilon^{BP}(t))^T,$$

$$\mathbf{C}_1=\begin{bmatrix} 0.793 & -0.076 & 0.000 & -0.001 \\ -0.035 & 0.590 & 0.098 & 0.047 \\ -0.044 & 0.014 & 0.755 & 0.013 \\ -0.022 & -0.224 & 0.063 & 0.856 \end{bmatrix},$$

$$\mathbf{C}_2=\begin{bmatrix} -0.041 & 0.029 & 0.000 & 0.026 \\ -0.001 & 0.085 & -0.079 & 0.030 \\ 0.010 & -0.022 & -0.079 & 0.072 \\ -0.020 & 0.024 & -0.016 & 0.066 \end{bmatrix},$$

$$\mathbf{C}_3=\begin{bmatrix} 0.003 & 0.081 & -0.058 & -0.037 \\ 0.091 & 0.113 & -0.163 & -0.044 \\ 0.046 & 0.099 & -0.099 & -0.051 \\ 0.075 & 0.302 & -0.269 & -0.116 \end{bmatrix},$$

$$\mathbf{C}_4=\begin{bmatrix} -0.038 & -0.089 & 0.081 & -0.002 \\ -0.043 & -0.248 & 0.220 & 0.023 \\ 0.008 & -0.196 & 0.142 & 0.016 \\ -0.019 & -0.305 & 0.282 & -0.014 \end{bmatrix},$$

Table 2.5 give error variances and AIC values for this fit.

In the MAR(4) model (2.5), each exchange rate has been decomposed and the irregulars of the four currencies have been fitted with a Vector Autoregressive model. The autoregressive model explains the relationship between the present irregulars and the past irregulars. Next we check the mutual contemporary relationship between the four final residuals of the vector autoregressive fit using principal component analysis. The results are shown in Table 2.6.

We see that while all the four currencies have a common factor (first principal component), the JY residual is rather isolated (by the second principal component) from the three European currencies; and among the three European currencies the BP residual is rather distant (by the third and the fourth principal components) from GM and SF. The residuals of SF and GM have mostly common factors; except the fourth principal component whose

TABLE 2.5.

	Error Variances			
	JY	GM	SF	BP
JY	0.198	0.130	0.143	0.117
GM	0.130	0.298	0.298	0.255
SF	0.143	0.298	0.372	0.266
BP	0.117	0.255	0.266	0.381

Order	0	1	2	3	4	5
AIC	3636.35	7.97	2.52	5.98	0.00	3.06
Order	6	7	8	9	10	11
AIC	11.84	6.61	6.26	31.93	25.88	28.33

Prediction error variances and AIC for the order 4 autoregressive fit to the residuals $\hat{\mathbf{I}}(t)$ of the (91,7) symmetric decomposition. The AIC has been adjusted by subtracting its minimum value.

TABLE 2.6.

Principal Component		1	2	3	4
Eigenvalue		2.956	0.629	0.315	0.100
Contribution		73.89%	15.73%	7.87%	2.51%
Cumulative Contribution		73.89%	89.62%	97.49%	100.00%
Eigenvectors	JY	0.404	0.901	0.160	0.012
	GM	0.549	−0.179	−0.319	−0.752
	SF	0.539	−0.159	−0.512	0.649
	BF	0.495	−0.362	0.781	0.117

Principal components analysis for the residual $\mathrm{e}(t)$ of the order 4 vector autoregressive fit to the residuals $\hat{\mathbf{I}}(t)$ of the (91,7) symmetric decomposition.

explanatory power is, however, very small (2.5%).

The coefficient matrices in (2.5) suggest an univariate autoregressive fit of order one for JY. Using S-PLUS, this leads to the fitted model

$$\hat{I}(t)=0.724\,\hat{I}(t-1)+\epsilon_1(t)$$

with prediction error variance 0.200 as compared to 0.198 for the MAR(4) model fit. This suggests that a good fit to the JY series is the parsimonious model

$$\begin{aligned}Y(t)&=L(t)+S(t)+0.724I(t-1)+\epsilon_1(t)\\&=0.724Y(t-1)+[L(t)-0.724L(t-1)]+[S(t)-0.724S(t-1)]+\epsilon_1(t).\end{aligned}$$

III. *Decomposition of Time Series Using Lagged Values*

(i) Methodology of Predictive Smoothing

The methods in the preceding section provide an ex-post analysis of a complete time series. That is, it provides a description in terms of a decomposition of how the series developed over time by using at time t both values prior and posterior to t. However, we may also want to decompose the time series today, at time t, using all the data available up to time t. This leads to a decomposition of the time series $Y(t)$ into what we call predictive long- and short-term trends as well as an irregular term. They are defined as

$$\begin{aligned}L_P(t)&=\sum_{k=t-TL}^{t} w_k(t)\mu(k),\\S_P(t)&=\sum_{k=t-TS}^{t} v_k(t)[\mu(k)-L_P(k)],\\I_P(t)&=Y(t)-[L_P(t)+S_P(t)]\end{aligned}\tag{3.1}$$

where $w_k(t)$ and $v_k(t)$ are weights. For instance, we might use

$$w_k(t)=\frac{\lambda_1^{t-k}}{\sum_{k=1}^{t}\lambda_1^{t-k}},\ 0<\lambda_1<1 \text{ and } v_k(t)=\frac{\lambda_2^{t-k}}{\sum_{k=1}^{t}\lambda_2^{t-k}},\ 0<\lambda_2<1,\ k=1,\ \ldots,\ t$$

with λ_1 close to one, such as 0.94; and λ_2 closer to zero, such as 0.5 or smaller. These are the ARIMA (0, 1, 1) weights used by J. P. Morgan's RiskMetrics^{TM} (1994), except for the standardization constant in the denominator.

The empirical versions $\hat{L}_P(t)$, $\hat{S}_P(t)$, and $\hat{I}_P(t)$ of (3.1) are obtained by replacing $\mu(k)$ by $Y(k)$ in the formula (3.1).

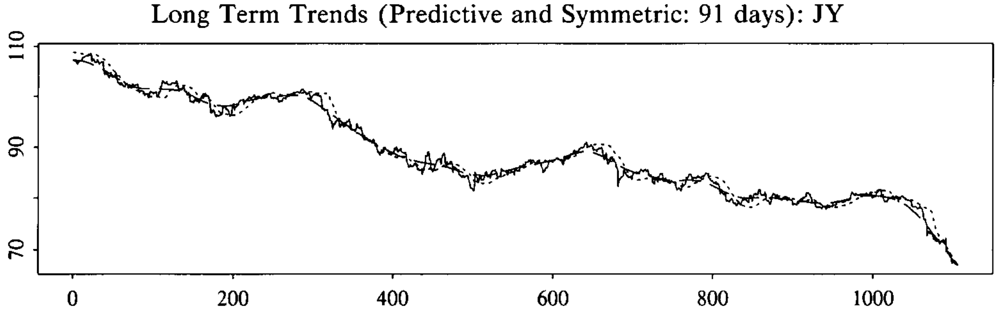
In the case that only lagged values can be used in the modeling of the series, uniform moving averages have serious biases over time periods with monotone trends and the locally linear method described in Section II is superior in this case. See Section VI for an analysis of this bias.

(ii) Long- and Short-Term Trends and the Residuals in Predictive Decompositions

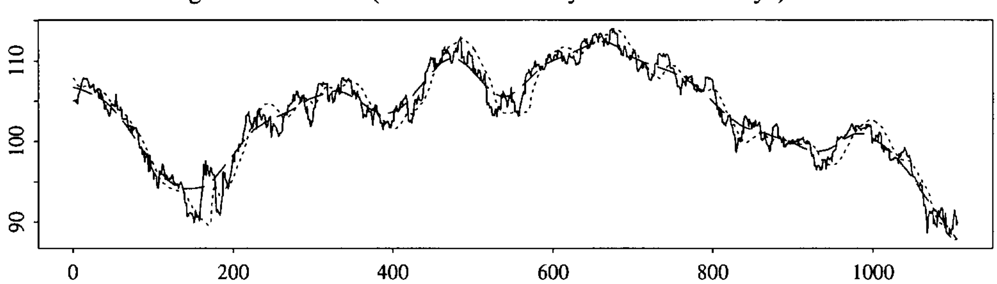
Figure 3.1 compares the past (ex-post, symmetric) and predictive long-term trends $\hat{L}(t)$

FIGURE 3.1.

Long Term Trends (Predictive and Symmetric: 91 days): JY



Long Term Trends (Predictive and Symmetric: 91 days): GM



The first graph gives the indexed currency exchange rate of JY —— and the symmetric – – – and predictive ········ 91-day long-term trends. The second graph shows the same for GM.

TABLE 3.1.

	91 days JY	GM	SF	BP	181 days JY	GM	SF	BP
JY	1.00	0.07	0.46	−0.46	1.00	−0.05	0.31	−0.30
GM	0.07	1.00	0.87	0.64	−0.05	1.00	0.87	0.78
SF	0.46	0.87	1.00	0.41	0.31	0.87	1.00	0.66
BP	−0.46	0.64	0.41	1.00	−0.30	0.78	0.66	1.00

Correlations of long-term predictive trends for the four currencies.

and $\hat{L}_P(t)$, based on the locally linear method (weights), for the JY and GM series. The graphs show how the predictive smoother $\hat{L}_P(t)$ overshoots the real series when there are sudden sustained drops in exchange rates while it undershoots in the case of sudden sustained increases. It illustrates the perils of prediction. We also computed $\hat{L}(t)$ and $\hat{L}_P(t)$ based on uniform moving average weights. This $\hat{L}_P(t)$ (not shown here) overshoots and undershoots the "true" trend $\hat{L}(t)$ much more than the locally linear $\hat{L}_P(t)$.

Table 3.1 shows the correlation coefficients of $L_P(t)$ of the four currencies. We see that the correlation between JY and GM is very small, whereas the correlation between JY and SF is 0.46 in the case of 91-day trend. The correlation of GM and SF is very high, 0.87. These results are consistent with those based on symmetric smoothing noted in Section II.

We graphed the predictive short-term trends, residual series, and long-term trends for the four exchange rates when the long- and short-term trends are based on 91- and 7-day locally linear smoothers. The long-term prediction starts 91 days after the beginning (January 5, 1995) of the exchange rate series while the short-term and residual series start 91+7=98 days after January 5, 1992. All the predictive series stop on April 14, 1995, one day before the last value in the series. The results (not shown here) were very similar to the symmetric decompositions of Section II, except the long- and short-term trends in the predictive decomposition are more variable.

We check the correlations between $S_P(t)$ for the four currencies in Table 3.2. Again the correlations are quite similar to those of the symmetric decompositions. Table 3.3 shows correlation coefficients between $I_P(t)$ for the four currencies and shows that the correlations are smaller (by about 0.1) than those in Table 2.3 in the previous section. These are the effects of not using the data posterior to the time t.

Table 3.4 shows that the predictive decomposition is very successful in achieving orthogonality between the terms in the decomposition for JY and is also quite successful for the other three European currencies, except the remaining small correlations between $L_P(t)$

TABLE 3.2.

	(91,7) days				(181,15) days			
	JY	GM	SF	BP	JY	GM	SF	BP
JY	1.00	0.41	0.43	0.11	1.00	0.40	0.42	0.08
GM	0.41	1.00	0.91	0.66	0.40	1.00	0.94	0.74
SF	0.43	0.91	1.00	0.60	0.42	0.94	1.00	0.69
BP	0.11	0.66	0.60	1.00	0.08	0.74	0.69	1.00

Correlations between short-term predictive trends for the four currencies.

TABLE 3.3.

	(91,7) days				(181,15) days			
	JY	GM	SF	BP	JY	GM	SF	BP
JY	1.00	0.37	0.37	0.27	1.00	0.41	0.37	0.31
GM	0.37	1.00	0.80	0.61	0.41	1.00	0.89	0.70
SF	0.37	0.80	1.00	0.59	0.37	0.89	1.00	0.66
BP	0.27	0.61	0.59	1.00	0.31	0.70	0.66	1.00

Correlations between residuals for the four currencies.

TABLE 3.4.

Correlation for (91,7) day trends	JAPANESE YEN				GERMAN MARKS			
	$\hat{L}_P(t)$	$\hat{S}_P(t)$	$\hat{L}_P(t)+\hat{S}_P(t)$	$\hat{I}_P(t)$	$\hat{L}_P(t)$	$\hat{S}_P(t)$	$\hat{L}_P(t)+\hat{S}_P(t)$	$\hat{I}_P(t)$
$\hat{L}_P(t)$	1.00	−0.09	0.99	−0.03	1.00	−0.26	0.96	−0.01
$\hat{S}_P(t)$	−0.09	1.00	0.05	−0.10	−0.26	1.00	0.01	−0.13
$\hat{L}_P(t)+\hat{S}_P(t)$	0.99	0.05	1.00	−0.04	0.96	0.01	1.00	−0.04
$\hat{I}_P(t)$	−0.03	−0.10	−0.04	1.00	−0.01	−0.13	−0.04	1.00
Correlation for (91,7) day trends	SWISS FRANCS				BRITISH POUNDS			
	$\hat{L}_P(t)$	$\hat{S}_P(t)$	$\hat{L}_P(t)+\hat{S}_P(t)$	$\hat{I}_P(t)$	$\hat{L}_P(t)$	$\hat{S}_P(t)$	$\hat{L}_P(t)+\hat{S}_P(t)$	$\hat{I}_P(t)$
$\hat{L}_P(t)$	1.00	−0.23	0.97	0.00	1.00	−0.32	0.97	0.00
$\hat{S}_P(t)$	−0.23	1.00	0.03	−0.11	−0.32	1.00	−0.07	−0.17
$\hat{L}_P(t)+\hat{S}_P(t)$	0.97	0.03	1.00	−0.03	0.97	−0.07	1.00	−0.05
$\hat{I}_P(t)$	0.00	−0.11	−0.03	1.00	0.00	−0.17	−0.05	1.00

Correlation for (181,15) day trends	JAPANESE YEN				GERMAN MARKS			
	$\hat{L}_P(t)$	$\hat{S}_P(t)$	$\hat{L}_P(t)+\hat{S}_P(t)$	$\hat{I}_P(t)$	$\hat{L}_P(t)$	$\hat{S}_P(t)$	$\hat{L}_P(t)+\hat{S}_P(t)$	$\hat{I}_P(t)$
$\hat{L}_P(t)$	1.00	0.03	0.97	−0.01	1.00	−0.36	0.90	0.00
$\hat{S}_P(t)$	0.03	1.00	0.28	−0.20	−0.36	1.00	0.08	−0.22
$\hat{L}_P(t)+\hat{S}_P(t)$	0.97	0.28	1.00	−0.06	0.90	0.08	1.00	−0.10
$\hat{I}_P(t)$	−0.01	−0.20	−0.06	1.00	0.00	−0.22	−0.10	1.00
Correlation for (181,15) day trends	SWISS FRANCS				BRITISH POUNDS			
	$\hat{L}_P(t)$	$\hat{S}_P(t)$	$\hat{L}_P(t)+\hat{S}_P(t)$	$\hat{I}_P(t)$	$\hat{L}_P(t)$	$\hat{S}_P(t)$	$\hat{L}_P(t)+\hat{S}_P(t)$	$\hat{I}_P(t)$
$\hat{L}_P(t)$	1.00	−0.28	0.91	0.01	1.00	−0.57	0.90	0.00
$\hat{S}_P(t)$	−0.28	1.00	0.15	−0.21	−0.57	1.00	−0.15	−0.22
$\hat{L}_P(t)+\hat{S}_P(t)$	0.91	0.15	1.00	−0.08	0.90	−0.15	1.00	−0.11
$\hat{I}_P(t)$	0.01	−0.21	−0.08	1.00	0.00	−0.22	−0.11	1.00

Correlations between components in the predictive decompositions.

and $S_P(t)$.

(iii) Autoregressive Analysis of the Predictive Irregular Series

Table 3.5 shows the autoregressive fit to the predictive residuals $\hat{I}_P(t)$. The Akaike Information Criteria applied to the four variate residuals selects order 3.

The four-variate autoregressive fitted model is:

$$\mathbf{I}_P(t)=\mathbf{C}_1\cdot\mathbf{I}_P(t-1)+\mathbf{C}_2\cdot\mathbf{I}_P(t-2)+\mathbf{C}_3\cdot\mathbf{I}_P(t-3)+\mathbf{e}_P(t), \tag{3.2}$$

where

$$\mathbf{I}_P(t)=(I_P^{JY}(t),\ I_P^{GM}(t),\ I_P^{SF}(t),\ I_P^{BP}(t))^T,$$
$$\mathbf{e}_P(t)=(\epsilon_P^{JY}(t),\ \epsilon_P^{GM}(t),\ \epsilon_P^{SF}(t),\ \epsilon_P^{BP}(t))^T,$$

$$\mathbf{C}_1=\begin{bmatrix} 0.149 & -0.061 & 0.008 & -0.046 \\ -0.024 & -0.010 & 0.055 & 0.007 \\ -0.034 & -0.036 & 0.176 & -0.036 \\ -0.014 & -0.051 & -0.078 & 0.188 \end{bmatrix},$$

$$\mathbf{C}_2=\begin{bmatrix} -0.183 & -0.069 & 0.041 & 0.020 \\ -0.008 & -0.110 & -0.056 & -0.014 \\ -0.000 & -0.075 & -0.173 & 0.029 \\ -0.022 & -0.154 & 0.075 & -0.116 \end{bmatrix},$$

$$\mathbf{C}_3=\begin{bmatrix} -0.110 & 0.066 & -0.098 & -0.011 \\ 0.012 & -0.043 & -0.150 & 0.004 \\ -0.024 & 0.058 & -0.238 & 0.028 \\ 0.016 & 0.047 & -0.142 & -0.078 \end{bmatrix},$$

The coefficients in (3.2) suggest an univariate order 2 autoregressive fit to the JY exchange rate series. Using S-PLUS, this fit is

$$Y(t)=\hat{L}_P(t)+\hat{S}_P(t)+0.137\,\hat{I}_P(t-1)-0.197\,\hat{I}_P(t-2)+\hat{\epsilon}_P(t), \tag{3.3}$$

with residual variance 0.076, which should be compared with the variance 0.080 of $\hat{I}_P(t)$. This shows that the decomposition $Y(t)=\hat{L}_P(t)+\hat{S}_P(t)+$ i.i.d. noise is very successful in that the

TABLE 3.5.

	Error Variances			
	JY	GM	SF	BP
JY	0.073	0.028	0.033	0.026
GM	0.028	0.081	0.074	0.062
SF	0.028	0.074	0.109	0.072
BP	0.026	0.062	0.072	0.131

Order	0	1	2	3	4	5
AIC	241.90	175.14	53.57	0.00	16.67	19.83

Prediction error variances and AIC for the order 3 autoregressive fit to the residuals $\mathbf{I}_P(t)$ of the (91,7) predictive decomposition. The AIC has been adjusted by subtracting its smallest value.

TABLE 3.6.

Principal Component		1	2	3	4
Eigenvalue		2.548	0.791	0.447	0.214
Contribution		63.70%	19.78%	11.18%	5.34%
Cumulative Contribution		63.70%	83.48%	94.66%	100.00%
Eigenvectors	JY	0.351	0.925	0.147	0.001
	GM	0.562	−0.149	−0.395	−0.711
	SF	0.561	−0.149	−0.411	0.703
	BF	0.497	−0.317	0.808	0.010

Principal components analysis for the residuals of the order 3 vector autoregressive fit to the residuals $\mathbf{I}_P(t)$ of (91,7) predictive decomposition.

reduction in variance by going to the more complex model (3.3) is only $0.080-0.076=0.004$.

Remark 3.1. The (91,7) predictive decomposition uses 91 days prior to t to make the long-term prediction $L_P(t)$ and seven days prior to t to make the short-term prediction $S_P(t)$. This short-term predictor can be regarded as a "local autoregressive" fit. It is a weighted linear fit using seven lagged values. Autoregressive techniques applies to the residuals $I_P(t)$ are teasing out possible overall (nonlocal) linear dependencies on lagged values. For instance, with JY there were slight autoregressive dependencies of $I_P(t)$ on the two first lagged values.

Remark 3.2. For a discussion of conditions needed for the existence of nonparametric time-series models of the type considered in this paper, see Masry and Tjóstheim (1995).

Remark 3.3. Since weekend trades are fewer than weekday trades and subject to different rules than weekday trades, we investigated a possible weekend effect by redoing the analysis using only weekdays for the (91,7) decomposition. The results (not shown here) did change a little. The Akaike Information Criteria selects an order 6 multivariate autoregressive fit with the lag one through three coefficients explains most of the variability with one interesting exception: The lag 6 (weekly) coefficient for BP is large for GM and SF. The multivariate lag one coefficient for JY is 0.149 when weekends are included and 0.125 when weekends are excluded.

IV. *Implications for Risk Measurement*

Based on the decompositions in the previous sections, we briefly describe a few possible applications to risk measurement. Recall that the decomposed components have very small correlations. We utilize these properties within a series and also their relationship with decomposed components of other series.

The basic idea for the applications is based on the interpretation that long- and short-term trends represent macro and micro aspects of the time-series movements, respectively; that the irregular part represents daily stochastics; and that the decomposition decomposes the time-series movements into nearly orthogonal terms. This interpretation for three components, i.e., three levels of uncertainties, is heuristic and depends on the time span used in the smoothing procedure. In order to confirm it we need to do a further study to relate the components with macro- and micro-economic indices of the countries. So the terms, macro, micro, and daily moves used in the following suggestions for applications are based on our heuristic interpretations.

Table 4.1 and 4.2 show the matrices of correlation coefficients of daily increments of the decomposed time-series components, $L(t)$ and $S(t)$, and the final AR-residuals $\epsilon_{AR}(t)$, for the four currencies. They are shown for both the symmetric and predictive smoothing methods, respectively.

From Box-Plots (not shown here) for daily increments $Y(t), L(t), S(t), I(t)$, and the final AR residuals, we conclude that the relative magnitude of the increments of $L(t)$ and $S(t)$ are small compared to that of $I(t)$ in the symmetric smoothing case; but they are not so small, especially $S(t)$ is not small, in the predictive smoothing case. To be precise, for JY during our data period, $dL(t), dS(t)$, and AR-residual of $I(t)$ take values mostly between -0.10 and 0.05,

TABLE 4.1.

	$dL(t)$ JY	$dL(t)$ GM	$dL(t)$ SF	$dL(t)$ BP	$dS(t)$ JY	$dS(t)$ GM	$dS(t)$ SF	$dS(t)$ BP	*ARres* JY	*ARres* GM	*ARres* SF	*ARres* BP
$dL(t)$: *JY*	1.00	0.56	0.61	0.29	0.10	0.00	0.01	0.00	−0.02	0.01	0.01	−0.01
$dL(t)$: *GM*	0.56	1.00	0.95	0.72	0.00	0.05	0.06	0.09	0.02	−0.02	−0.02	−0.06
$dL(t)$: *SF*	0.61	0.95	1.00	0.73	0.01	0.06	0.08	0.10	0.01	0.00	0.00	−0.05
$dL(t)$: *BP*	0.29	0.72	0.73	1.00	−0.03	0.04	0.05	0.11	0.00	−0.02	−0.01	−0.05
$dS(t)$: *JY*	0.10	0.00	0.01	−0.03	1.00	0.54	0.51	0.27	0.06	0.03	0.04	0.01
$dS(t)$: *GM*	0.00	0.05	0.06	0.04	0.54	1.00	0.93	0.68	0.02	0.03	0.04	0.02
$dS(t)$: *SF*	0.01	0.06	0.08	0.05	0.51	0.93	1.00	0.69	0.03	0.02	0.03	0.01
$dS(t)$: *BP*	0.00	0.09	0.10	0.11	0.27	0.68	0.69	1.00	0.04	0.00	−0.01	0.05
ARres: *JY*	−0.02	0.02	0.01	0.00	0.06	0.02	0.03	0.04	1.00	0.54	0.53	0.42
ARres: *GM*	0.01	−0.02	0.00	−0.02	0.03	0.03	0.02	0.00	0.54	1.00	0.90	0.76
ARres: *SF*	0.01	−0.02	0.00	−0.01	0.04	0.04	0.03	−0.01	0.53	0.90	1.00	0.71
ARres: *BF*	−0.01	−0.06	−0.05	−0.05	0.01	0.02	0.01	0.05	0.42	0.76	0.71	1.00

Correlations between $dL(t)$, $dS(t)$ and *AR*-residuals for the symmetric decomposition (91,7).

TABLE 4.2.

	$dL_P(t)$ JY	$dL_P(t)$ GM	$dL_P(t)$ SF	$dL_P(t)$ BP	$dS_P(t)$ JY	$dS_P(t)$ GM	$dS_P(t)$ SF	$dS_P(t)$ BP	*ARres* JY	*ARres* GM	*ARres* SF	*ARres* BP
$dL_P(t)$: *JY*	1.00	0.33	0.41	0.09	−0.12	0.00	−0.03	0.03	0.00	0.02	0.06	0.03
$dL_P(t)$: *GM*	0.33	1.00	0.85	0.37	−0.04	−0.10	−0.05	−0.04	0.00	−0.05	0.01	−0.01
$dL_P(t)$: *SF*	0.41	0.85	1.00	0.43	−0.04	−0.05	−0.09	−0.02	−0.02	−0.04	−0.02	−0.01
$dL_P(t)$: *BP*	0.09	0.37	0.43	1.00	0.03	0.01	−0.01	−0.21	−0.01	0.02	−0.01	−0.06
$dS_P(t)$: *JY*	−0.12	−0.04	−0.04	0.03	1.00	0.39	0.40	0.23	0.31	0.30	0.28	0.21
$dS_P(t)$: *GM*	0.00	−0.10	−0.05	0.01	0.39	1.00	0.82	0.60	0.25	0.45	0.47	0.39
$dS_P(t)$: *SF*	−0.03	−0.05	−0.09	−0.01	0.40	0.82	1.00	0.56	0.22	0.44	0.45	0.33
$dS_P(t)$: *BP*	0.03	−0.04	−0.02	−0.21	0.23	0.60	0.56	1.00	0.19	0.36	0.37	0.36
ARres: *JY*	0.00	0.00	−0.02	−0.01	0.31	0.25	0.22	0.19	1.00	0.37	0.37	0.27
ARres: *GM*	0.02	−0.05	−0.04	0.02	0.30	0.45	0.44	0.36	0.37	1.00	0.79	0.60
ARres: *SF*	0.06	0.01	−0.02	−0.01	0.28	0.47	0.45	0.37	0.37	0.79	1.00	0.60
ARres: *BF*	0.03	−0.01	−0.01	−0.06	0.21	0.39	0.33	0.36	0.27	0.60	0.60	1.00

Correlations between $dL_P(t)$, $dS_P(t)$ and *AR*-residuals for the predictive decomposition (91,7).

TABLE 4.3.

(a) Symmetric (91,7)

	JY		GM		SF		BP	
	10%	90%	10%	90%	10%	90%	10%	90%
$dL(t)$	−0.101	0.050	−0.122	0.098	−0.141	0.112	−0.089	0.120
$dS(t)$	−0.121	0.118	−0.173	0.206	−0.198	0.223	−0.235	0.199
$\epsilon_{AR}(t)$	−0.495	0.448	−0.658	0.629	−0.699	0.692	−0.718	0.668

(b) Predictive (91,7)

$dL_P(t)$	−0.186	0.089	−0.184	0.171	−0.223	0.204	−0.168	0.195
$dS_P(t)$	−0.386	0.415	−0.531	0.554	−0.606	0.604	−0.606	0.591
$\epsilon_{P,AR}(t)$	−0.194	0.202	−0.238	0.267	−0.282	0.296	−0.294	0.288

(a) 10% and 90% percentiles of $dL(t)$, $dS(t)$, and *AR*-residuals of four currencies based on the symmetric decompositions (91,7). (b) gives the same for the predictive (91,7) decomposition.

-0.12 and 0.12, -0.5 and 0.5, respectively, in our symmetric (91,7) decomposition. They are mostly between -0.19 and 0.09, -0.39 and 0.41, -0.19 and 0.20, respectively, in our predictive (91,7) decomposition. For other currencies see Table 4.3, where 10% and 90% percentiles of $dL(t)$, $dS(t)$, and AR-residual for each currency are shown. These results are important because they indicate the relative magnitudes of the contributions of $dL(t)$, $dS(t)$ and $dI(t)$ to the whole increment of $Y(t)$, that is to "Risk". The algorithm of LOWESS may be relevant to these.

Looking at time-series data retrospectively (with symmetric smoothing), we see how it behaved as results of market tradings; and looking at the uncertainties at each time points (with predictive smoothing), we simulate what the practitioners have been facing at each time point. The applications of our decompositions would utilize these two viewpoints.

(i) Behaviors of the Four Currencies Against U.S. Dollars in Three Levels

Our currency exchange rates data represent the relative weakness of each currency against U.S. dollars. The correlation coefficients of $dL(t)$ and $dS(t)$ of JY with those of GM are 0.56 and 0.54 in symmetric smoothing and they are 0.33 and 0.39 in predictive smoothing. These numbers are about the same for SF, but they are much smaller for BP. These numbers are smaller in the predictive smoothing case. Note that $dL(t)$ of JY and BP are uncorrelated in the predictive smoothing case though there is some correlation (0.29) in symmetric smoothing. A further study may be of interest to find what are the causes of this.

(ii) Calculating VaR (Value at Risk)

The Value at Risk (VaR) is a lower threshold v for $dY(t)=Y(t+1)-Y(t)$ defined by $P(dY(t)\leq v)=0.01$. It is linked to the standard deviation of $dY(t)$. Assuming that the long-term trend does not have any sudden changes, it should be possible to reduce the standard deviation of $dY(t)$. Note that $dY(t)=dL_P(t)+dS_P(t)+dI_P(t)$ and our idea is that $dL_P(t)$ may almost be a predictable constant under the assumption. See Figure 4.1.

At a time point t, our prediction $\hat{Y}(t+1)$ of $Y(t+1)$ based on data up to time t is

$$\hat{Y}(t+1)=L_P(t)+\{L_P(t)-L_P(t-1)\}+S_P(t)+(\text{time-dependent autoregressive form of } I_P(t)).$$

Then, under our assumption that $L_P(t)$ is locally constant, $dY(t)$ is decomposed as follows:

$$\begin{aligned} dY(t) &= Y(t+1)-Y(t) \\ &= \{Y(t+1)-\hat{Y}(t+1)\}+\{\hat{Y}(t+1)-Y(t)\} \\ &= dY_P(t)+dS_P(t)+(a_t-1)\,I_P(t)+b_t\,I_P(t-1)+\cdots \end{aligned}$$

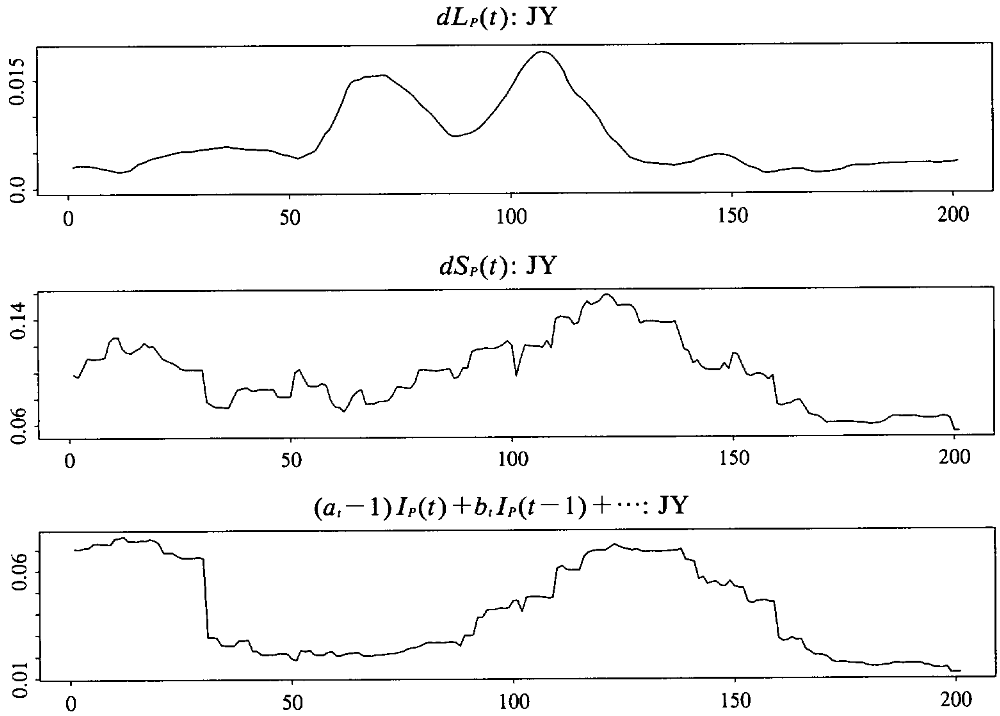
Regarding $dL_P(t)$ as a constant, the variance of $dY(t)$ is approximately the sum of the variances of $dS_P(t)$ and $I_P(t)$. See Figure 4.1. Following this intuition, we estimate VaR. Our estimate of the 1% percentile of $dY(t)$, or the Value at Risk, will be approximately

$$\text{VaR}=\{\hat{Y}(t+1)-Y(t)\}-c\hat{\sigma}$$

where $\hat{\sigma}$ stands for the estimated standard deviation of $[dS_P(t)+I_P(t)$ terms] and $-c$ is the 1% percentile of the distribution of $[\hat{Y}(t+1)-Y(t)]/\hat{\sigma}$.

A normal approximation would give $c=2.33$. In order to see how well our VaR works with this c, we calculated our VaR for 201 days in a row and compared it with two other

FIGURE 4.1.



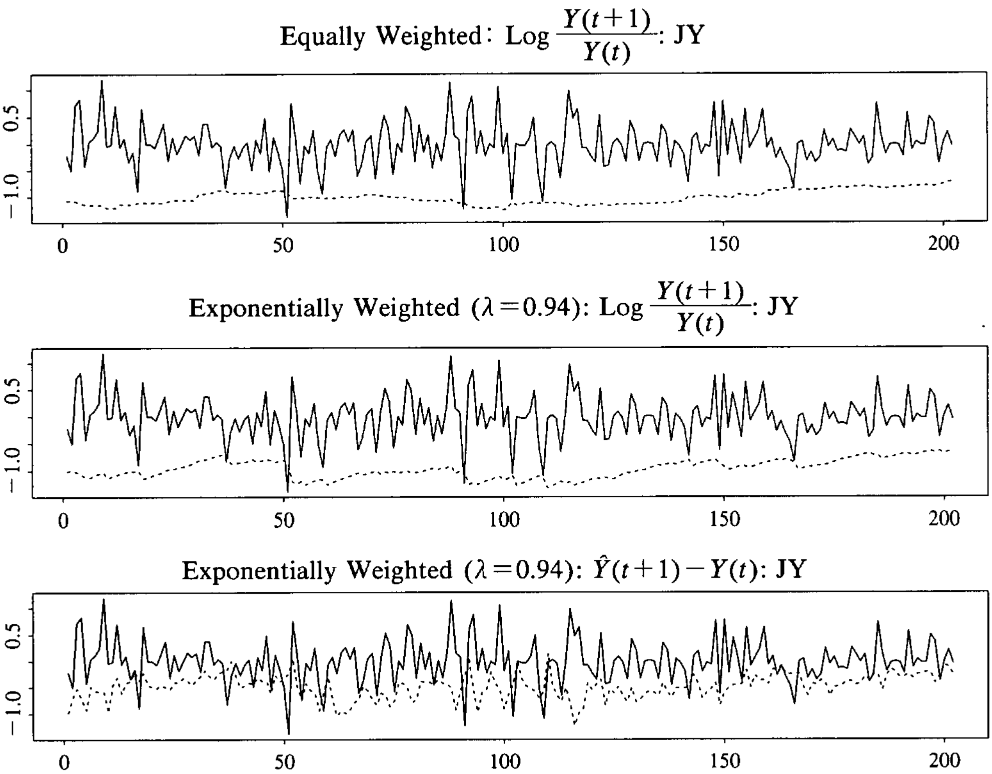
The first and the second graphs give daily increments of 91-day long-term trends and 91- and 7-day short-term trends for JY, respectively. The third one gives the I_P term.

standard VaR based on the equally weighted sample variance and J. P. Morgan's exponentially weighted sample variance. Our VaR with $c=2.33$ does not work well in the sense that the actual $dY(t)$ went below our VaR 22 times within 201 days while this happened three times for J. P. Morgan's method. See Figure 4.2. Our intuitive idea needs to use a more accurate critical constant c so that this frequency will be more compatible with 1% chance. One possibility is a bootstrap method. Our effort is to have a small and efficient $\hat{\sigma}$ by focusing on the practical uncertainties of VaR. Our $\hat{\sigma}$ is smaller than the standard errors of the other two methods in most cases.

(iii) Back Testing

Back testing of a VaR measurement system compares the time-series values of VaR calculated by the system under examination with the time-series values of realized values of the portfolio by counting how many times the realized values went over the estimated 1% percentile. Our two-step smoothing technique provides decompositions of the variability from two viewpoints. So it will be interesting to compare the time-series behavior of VaR with our decomposed $L(t)$, $S(t)$, and $I(t)$. When the realized value of a portfolio goes over VaR at a certain time, we may check which one of the three components is most responsible in a retrospective way (with symmetric decomposition), and we may also check what was known to the practitioners at the time in the predictive way. The same remarks apply to the

FIGURE 4.2.



Plots of actual daily increments of $Y(t)$ ——— and VaR -------- for JY. The first two figures are based on $\log[Y(t+1)/Y(t)]$ using equal weights and exponential weights with $\lambda=0.94$, respectively. The third figure is based on $\hat{Y}(t+1)-Y(t)$ using exponential weights with $\lambda=0.94$.

measurement of volatility in evaluating the increments of prices of derivatives. For these problems, it will also be interesting to compare the volatilities obtained from our predictive smoothing and the implied volatilities provided among the traders in the markets. We leave these interesting problems for later research.

V. *Volatility Measures and Their Time-Series Decompositions*

The predictive decomposition of Section III can be applied to time series other than original exchange rates. We next consider the problem of assessing the volatility of exchange rate series such as those considered in the earlier section.

(i) Volatility Measures

Many commonly used measures of volatility are based on the conditional variance $\sigma^2(t)$ given the past of the returns $X(t)=\log[Y(t)/Y(t-1)]$, e.g., ARCH (Engle, 1982) and GARCH (Bollerslev, 1986). The set of conditioning variables, which we denote by Ω_{t-1}, contains information on lagged values of $X(t)$ as well as on other exchange rates, interest rates, etc. in general, up to and including time $t-1$. Let $m(t)=E(X(t)|\Omega_{t-1})$ denote the conditional mean of the returns and let $V(t)=[X(t)-m(t)]^2$ denote the squared deviation from the conditional mean. By definition, the conditional variance $\sigma^2(t)$ given the past is $\sigma^2(t)=\mathrm{Var}(X(t)|\Omega_{t-1})=E[V(t)|\Omega_{t-1}]$. Thus, we can think of $\sigma^2(t)$ as the conditional mean of $V(t)$, and apply the predictive decomposition of Section III. To illustrate, we take Ω_{t-1} to be the set of lagged values $X(t-1)$, $X(t-2)$, ... of the returns. Because of the empirical evidence and computational convenience, $m(t)$ is often taken to be zero or a constant. We considered different candidates for $m(t)$ and found that the moving average $m(t)=(t-1)^{-1}\sum_{j=1}^{t-1}X(j)$ gave the best results in the sense of avoiding overfits. Using locally linear smoothers for $m(t)$ led to overfits, that is, they would tend to make $V(t)$ close to zero.

The long- and short-term trends in the conditional variance $\sigma^2(t)$ are

$$\sigma_L^2(t)=\sum_{k=1}^{t} w_k(t)E[V(k)],$$

$$\sigma_S^2(t)=\sum_{k=1}^{t} v_k(t)\{E[V(k)]-\sigma_L^2(k)\}.$$

Our assumption is that, to a close approximation, $I_P(t)=V(t)-[\sigma_L^2(t)+\sigma_S^2(t)]$ has conditional (given Ω_{t-1}) expected value zero (if not, a further decomposition of $I(t)$ as in Section III could be used).

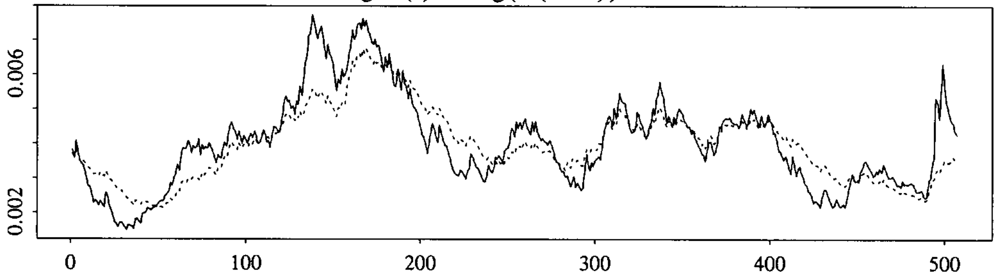
Before analyzing $V(t)$ for the exchange rate data, we removed the weekends. In the long- and short-term decomposition in the earlier sections it was convenient to leave them in since it makes the time axis correspond to calendar time and because this particular analysis is not sensitive to weekend effects. However, returns are sensitive to weekend effects, especially since the number of weekend transactions is a small fraction of the number of weekday transactions. This makes the absolute weekend returns very close to zero and the volatility measures biased toward zero. It would help to divide the returns by the logarithm of volume; however, this ratio would then have a much larger variance for weekends than for weekdays.

By using only weekday exchange rates, we will have returns only for Tuesday, Wednesday, Thursday, and Friday; that is for 689 days. The restriction to weekdays has the advantage of reducing the number of missing data days considerably. For JY there was only one missing data day while for GM and SF there were two and two missing data days. For these days we used the average of the returns of the two closest days where returns are available. For days with volume less than 200, where volume stands for the number of changes on the computer screen during a 24-hour period, we used the average of that days return with the returns of the two closest days where returns are available. The average volume (in number of changes per 24 hours) over weekdays for days with volume at least 200 was 3456 and exceptionally small volumes were rare.

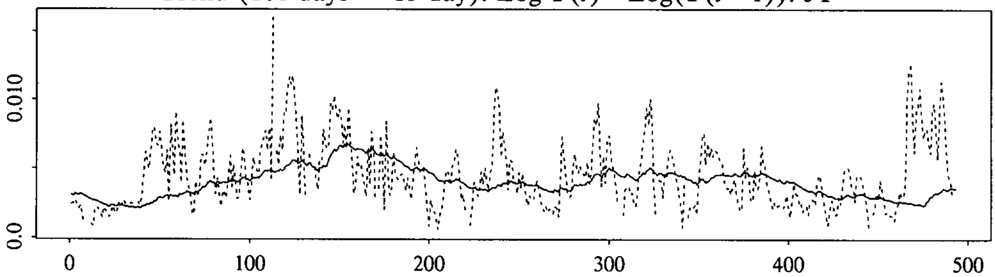
Figure 5.1 shows a graphical analysis of the volatility of JY. The first figure compares the 91- and 181-day long-term trends $\hat{\sigma}_L(t)$ and shows how the 91-day trend is sensitive to changes in $V(t)$ while the 181-day trend makes more moderate adjustments as $V(t)$ changes. The second one shows how the addition of the short-term trend fine tunes the long-term trend and results in a volatility measure which is very sensitive to the recent past.

Figure 5.2 gives a comparison of the volatilities of GM and SF. They resemble each other, but the SF are generally more volatile.

FIGURE 5.1. Standard Deviation Estimator: Long Term Trends (91 days and 181 days): Log $Y(t)-\text{Log}(Y(t-1))$: JY

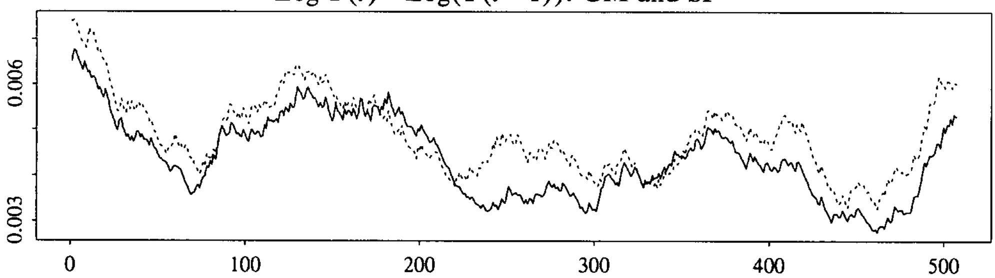


Standard Deviation Estimator: Long Term Trend and Long + Short Term Trend (181 days → 15 day): Log $Y(t)-\text{Log}(Y(t-1))$: JY

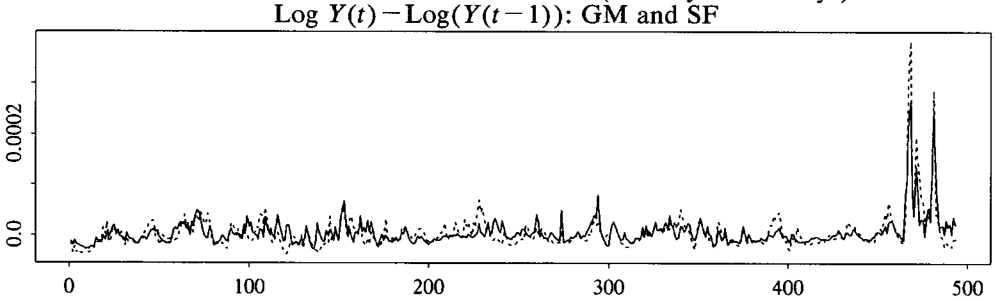


The first graph gives the long-term volatility of JY using predictive decomposition based on 91 ——, and 181 -------- day time intervals. The second graph gives the volatility of JY using (181,15) predictive decomposition. The solid line is the long-term trend while the dotted line is the sum of the long- and short-term trends.

FIGURE 5.2. Standard Deviation Estimator: Long Term Trends (181 days): Log $Y(t)-\text{Log}(Y(t-1))$: GM and SF



Variance Estimator: Short Term Trends (181 days → 15 days): Log $Y(t)-\text{Log}(Y(t-1))$: GM and SF



The first figure gives the long-term volatility of GM —— and SF --------. The estimators are 181-day predictive smoothers. The second one gives the short-term trends of GM —— and SF -------- based on the (181,15) predictive decomposition.

(ii) RiskMetrics^{TM}

The J.P.Morgan RiskMetrics^{TM} (1994) measure the volatility is an example of a predictive long-term trend kernel estimator. In the third edition of RiskMetrics^{TM} (page 80) they give

$$\hat{\sigma}^2_{JP}(t)=\sum_{k=-K}^{t} w(t-k)X^2(k)$$

with $w(t-k)=(1-\lambda)\lambda^{(t-k)}$, $\lambda=0.94$ and $X(0)$, $X(-1)$, ..., $X(-K)$ historical data. An approximation to $\hat{\sigma}^2_{JP}(t)$ based on the historical data is

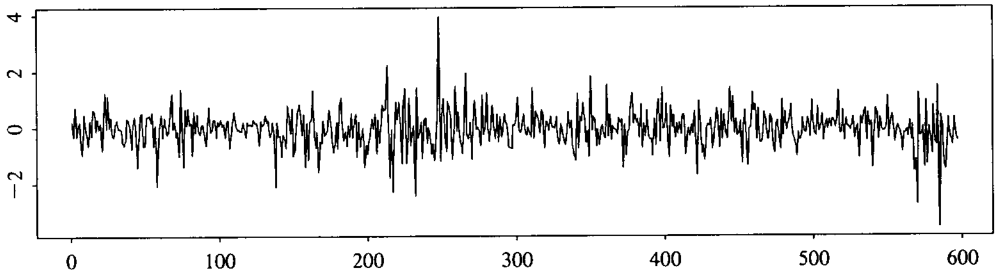
$$\tilde{\sigma}^2_{JP}(t)=\lambda\tilde{\sigma}^2_{JP}(t-1)+(1-\lambda)X^2(t)$$

where $\tilde{\sigma}^2_{JP}(0)=\frac{1}{K}\sum_{k=-K}^{0}X^2(k)$. This formula $\tilde{\sigma}^2_{JP}(t)$ is the one recommended by J. P. Morgan even though $\hat{\sigma}^2_{JP}(t)$ also can be computed for historical data and gives roughly the same results. However, $\tilde{\sigma}^2_{JP}(t)$ has the advantage of being easy to understand as an updating formula. Note that RiskMetrics^{TM} uses $m(t)=0$ and $V(t)=X^2(t)$. Other $m(t)$ were tried by J. P. Morgan in earlier versions.

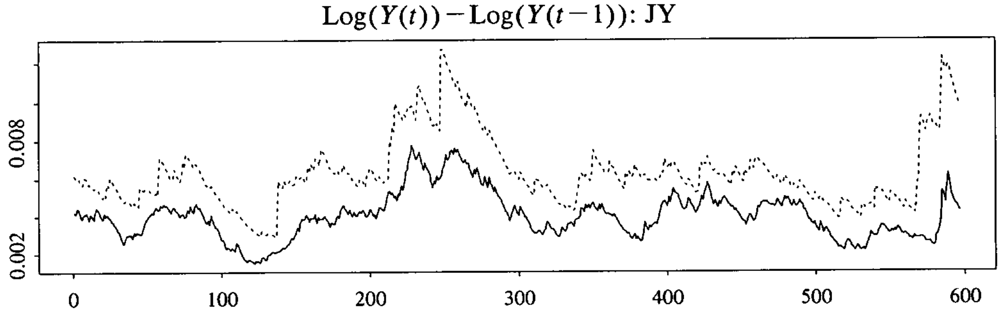
Figure 5.3 uses the Japanese exchange rate series to give a comparison between the square roots of the predictive 91-day long-term smoother and the RiskMetrics^{TM} exponential kernel estimator. Both are based on using the first 91 days as historical data. The figure shows that

FIGURE 5.3.

Percent of Changes: Log($Y(t)$)−Log($Y(t-1)$): JY



Standard Deviation Estimator: Predictive (91 days) and JPM (Approximation): Log($Y(t)$)−Log($Y(t-1)$): JY



The first figure gives the returns of the Japanese exchange rate series. The second figure gives the 91-day long-term trend volatility using our predictive decomposition —— and the J. P. Morgan's RiskMetrics^{TM} volatility -------- of JY using Tuesday-Friday data from the time period from January 6, 1992, to April 14, 1995.

both estimators give similar assessments of volatility with the 91-day predictive smoother having slightly less abrupt changes. In Section VI we argue that the locally linear predictive smoother shown in Figure 5.3 falls closer to the true volatility than RiskMetrics^{TM}. Note that RiskMetrics^{TM} is higher than $\hat{L}_P(t)$ because it is based on using $m(t) \equiv 0$ whereas $\hat{L}_P(t)$ is based on a data-based centering statistic $\hat{m}(t)$.

VI. *Asymptotic Analysis*

Asymptotic results and approximations that yield insights into our decompositions can be obtained by letting the number of terms T in the series $Y(t)$ tend to infinity. In particular, we can use approximations to quantify how sudden changes in the mean function of a time series is reflected in our long- and short-term trend functions.

The approximations also lead to ways of evaluating the performance of predictive smoothers by comparing them with symmetric smoothers. The reasoning is as follows: Because the symmetric smoothers are based on future as well as past values of a time series, they track the "true" trends in the time series more closely than the predictive smoothers. On the other hand, the predictive smoothers, such as J. P. Morgan's RiskMetrics^{TM}, have applications to volatility assessment which can be used in portfolio management. One way to choose among the many predictive smoothers available is to choose the one that falls closest to the symmetric smoother since the symmetric smoother tracks the "true" series better. Our approximations of this section show that in this sense the locally linear predictor is preferable to kernel estimators such as RiskMetrics^{TM} when $\mu'(t) \neq 0$.

We rescale by setting

$$u = \frac{k}{T+1}, \; v = \frac{t}{T+1}, \; h = \frac{M}{T}. \tag{6.1}$$

That is, the fixed time point t of interest is transformed to $v = t/(T+1)$, the summation index k is transformed to $u = k/(T+1)$, and the time interval $[-M, M]$ where the nonparametric regression smoother is nonzero is transformed to $(-h, h)$.

The notation is as in Section II. We define $\mu_h(u) = \mu([u(T+1)])$, $u \in (0, 1)$, where [ ] denotes the greatest integer function, and assume that $\mu_h(u) \to \mu_0(u)$ as $T \to \infty$, $u \in (0, 1)$, for some function $\mu_0(u)$ on $(0, 1)$ as $T \to \infty$. In other words, $\mu_0(u)$ is the limiting mean function of the time series $Y(t)$ after the transformation (6.1) of the time scale to (0,1).

(i) The Symmetric Regression Smoothers

In the case of the symmetric long-term smoothing kernel weights of Section II, we write $[-M_L, M_L]$ for the interval of time points k that the kernel is nonzero; and we set $h_L = M_L/T$. Using a very long-term smoother is equivalent to assuming that for some $0 < h_0 < 1$, $h_L \to h_0$ as $T \to \infty$. We let $L_h(u) = L([u(T+1)])$ be the long-term trend in the new units where L is based on the weights (2.1) and (2.2), then

Proposition 6.1. *Suppose the kernel K is a symmetric continuous density on* $[-1, 1]$ *and suppose that $\mu_0(u)$ is Riemann integrable. If $h = h_L \to h_0$ with $h_0 \in (0, 1)$ as $T \to \infty$, then, for each $v \in (h_0, 1-h_0)$,*

$$L_0(v) \equiv \lim_{T\to\infty} L_h(v) = h_0^{-1} \int_{v-h_0}^{v+h_0} \mu_0(u) K\left(\frac{u-v}{h_0}\right) du = \int_{-1}^{1} \mu_0(v+h_0 x) K(x) dx. \tag{6.2}$$

Proof. First apply the change of variable (6.1) to the sum (2.1) with weights (2.2). Next note that the limit of the sum is the Riemann integral (6.2).

We call the limit $L_0(v)$ in (6.2) the *asymptotic long-term trend.*

For the short-term kernel weights of Section II we introduce $h_S = M_S/L$ where $[-M_S, M_S]$ is the time interval where the kernel is nonzero. To use a short-term smoother is equivalent to assuming that $h_S \to 0$ as $T \to \infty$. We consider the following "large T" short-term trend for $v \in (h_0, 1-h_0)$:

Lemma 6.1. *Suppose that $\mu_0(u)$ and $L_0(u)$ have continuous third derivatives at v.* If $h = h_S \to 0$ *and $Th_S^3 \to \infty$ as $T \to \infty$, then*

$$\begin{aligned} S_h(v) &= (M_S)^{-1} \sum_u K\left(\frac{u-v}{h_S}\right) [\mu_0(u) - L_0(u)] \\ &= (h_S T)^{-1} \sum_u K\left(\frac{u-v}{h_S}\right) \Big[\mu_0(v) - L_0(v) + [\mu_0'(v) - L_0'(v)](u-v) \\ &\qquad + \frac{1}{2} [\mu_0''(v) - L_0''(v)](u-v)^2\Big] + O(h_S^3). \end{aligned} \tag{6.3}$$

Since the kernel K is symmetric, the first derivative term in the expansion (6.3) is zero. By setting $x = (u-v)/h_S$ and $\sigma_K^2 = \int_{-1}^{1} x^2 K(x) dx$ = the variance of K, we find the following approximation to the limiting short-term trend.

Proposition 6.2. *Under the conditions of Lemma* 6.1 *and Proposition* 6.1,

$$S_h(v) = \mu_0(v) - L_0(v) + \frac{1}{2} [\mu_0''(v) - L_0''(v)] h_S^2 \sigma_K^2 + O(h_S^3),\ h_0 < v < 1 - h_0. \tag{6.4}$$

This approximation shows how the short-term trend depends on the instantaneous value $\mu_0(v) - L_0(v)$, the second derivative, the bandwidth h_S, and the variance of K.

Next, we consider the case where h_L is of order between T^{-1} and h_S, that is $h_L \to 0$ and $(h_S/h_L) \to 0$. In this case the trend considered is moderately long. From (6.2) we find that the moderately long-term trend $L_h(v)$ with $h = h_L$ is

$$L_h(v) = \mu_0(v) + \frac{1}{2} \mu_0''(v) h_L^2 \sigma_K^2 + O(h_L^3) + O(h_S^2),\ h_0 < v < 1 - h_0. \tag{6.5}$$

By combining this with (6.4) we find

Proposition 6.3. *Suppose the conditions of Proposition* 6.1 *and Lemma* 6.1 *hold. If $h = h_L \to 0$ and $(h_S/h_L) \to 0$ as $T \to \infty$, then*

$$S_h(v) = -\frac{1}{2} \mu_0''(v) h_L^2 \sigma_K^2 + O(h_L^3),\ h_0 < v < 1 - h_0. \tag{6.6}$$

This approximation shows that the second derivative has a crucial influence on the short-term trend in the case of a "moderately long" long-term trend.

The idea of using h_L of different order corresponds to using different time spans or "different levels of uncertainty". See Miura and Kishino (1995).

(ii) The Predictive Regression Smoothers

We let $L_P^{(h)}(v)$ denote the long-term trend in the new units. If we apply the preceding limits and approximations to the predictive smothers in Section III based on weights of the form (2.2), we find the following approximations to L_P.

Proposition 6.4. *Under the conditions of Proposition* 6.1, *for each* $v \in \left(h_0, \frac{1}{2}\right)$,

$$L_P^{(0)}(v) \equiv \lim_{T \to \infty} L_P^{(h)}(v) = 2h_0^{-1} \int_{v-h_0}^{v} \mu_0(u) K\left(\frac{u-v}{h_0}\right) du = 2\int_{-1}^{0} \mu_0(v+h_0 x) K(x) dx. \tag{6.7}$$

We define the "large T" predictive short-term trend as

$$S_P^{(h)}(v) = \left(\frac{1}{2} M_S\right)^{-1} \sum_{u \le \frac{1}{2}} K\left(\frac{u-v}{h_S}\right) [\mu_0(u) - L_P^{(0)}(u)].$$

By Taylor expansion, as $T \to \infty$, $h = h_S \to 0$,

$$S_P^{(h)}(v) = \mu_0(v) - L_P^{(0)}(v) + [\mu_0'(v) - L_P^{(0)\prime}(v)] h_S \mu_{PK} + \frac{1}{2} [\mu_0''(v) - L_P^{(0)\prime\prime}(v)] h_S^2 \mu_{PK}^2 + O(h_S^3) \tag{6.8}$$

where $\mu_{PK} = 2\int_{-1}^{0} xK(x)dx$, $\sigma_{PK}^2 = 4\int_{-1}^{0} x^2 K(x)dx$, and $v \in \left(h_0, \frac{1}{2}\right)$.

We now note that the short-term trend depends crucially on the first derivative $\mu_0'(v)$ of the mean. That is, the short-term kernel regression smoother is very sensitive to sudden changes in the theoretical means $\{\mu(t)\}$. On the other hand, the locally linear smoothers discussed in Section II do not have this property. Using the results of Fan and Gijbels (1996) and some manipulations, it is possible to show that for the locally linear short-term predictive smoother defined by (2.3) and (2.4), the "large T" short-term trend is,

$$S_{L(P)}^{(h)}(v) = \mu_0(v) - L_P^{(0)}(v) + \frac{1}{2} [\mu_0''(v) - L_P^{(0)\prime\prime}(v)] h_S^2 \mu_{PK}^2 + O(h_S^3). \tag{6.9}$$

From this it follows that when $\mu_0'(v) \neq 0$ the locally linear predictive smoothers are much closer to the symmetric smoothers; and, in this sense, they are better than the kernel smoothers at tracking the "true" trends in the time series.

If we consider moderately long-term trends where $h_L \to 0$ and $(h_S/h_L) \to 0$, we obtain, for smoothers based on the weights (2.2), the approximations

$$L_P^{(h)}(v) = \mu_0(v) + \mu_0'(v) h_L \mu_{PK} + \frac{1}{2} \mu_0''(v) h_L^2 \mu_{PK}^2 + O(h_L^3). \tag{6.10}$$

$$S_P^{(h)}(v) = \mu_0'(v) h_L \mu_{PK} + O(h_L^2). \tag{6.11}$$

These approximations show more clearly the sensitivity of the predictive kernel regression smoothers to the first derivative $\mu_0'(v)$. The locally linear predictive smoothers do not share this sensitivity and are much closer to the symmetric linear smoothers.

References

Becker, R.A., J.M. Chambers, and A.R. Wilks, "The New *S* Languages," Wadsworth, California, 1988.

Beveridge, Stephen, and Charles R. Nelson, "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycles'," *Journal of Monetary Economics*, Vol. 7, No. 2, 1981, pp. 151-174.

Bollerslev, T., "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, Vol. 31, 1986, pp. 307-327.

Box, George E.P., and Gwilym M. Jenkins, "Time Series Analysis: Forecasting and Control," Oakland CA: Holden-Day, 1976.

Brock, William A., David A. Hsieh, and Blake LeBaron, "Nonlinear Dynamics, Chaos, and Instability," Cambridge, MA: The MIT Press, 1991.

Brock, W., W. Dechert, and J. Scheinkman, "A Test for Independence Based on the Correlation Dimension," Technical Report, University of Wisconsin, Madison, 1987.

Cleveland, W.S., "Robusts Locally Weighted Regression and Smoothing Scatterplot," *Journal of American Statistical Association*, 74, 1979, pp. 829-836.

Cleveland, W.S., S.J. Devlin, D.R. Schapira, and I.J. Terpenning, "The SABL Seasonal and Calendar Adjustment Package," Computing Information Library, Bell Laboratories, 1981.

Cleveland, W.S., and S.J. Devlin, "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting," *Journal of American Statistical Association*, 83, 1988, pp. 596-610.

Doksum Kjell, Ryozo Miura, and Hiroaki Yamauchi, "On Financial Time Series Decompositions with Applications to Volatility," IMES Discussion Paper Series No.98-E-6, IMES, Bank of Japan, 1998.

Doksum, Kjell, and Alex Samarov, "Nonparametric Estimation of Global Functionals and a Measure of the Explanatory Power of Covariates in Regression," *The Annals of Statistics*, 23, 1995, pp. 1443-1473.

Enders, Walter, "Applied Economic Time Series," New York, NY: John Wiley & Sons, Inc., 1994.

Engle, Robert E., "Autoregressive Conditional Heteroskedasticity," *Econometrica*, Vol. 50, 1982, pp. 987-1007.

Fan, Jianqing, and Irène Gijbels, "Local Polynomial Modeling and Its Applications," London, England: Chapman and Hall, 1996.

Fan, Jianqing, "Local Linear Regression Smoothers and Their Minimax Efficiency," *Annals of Statistics*, Vol. 21, 1993, pp. 196-216.

Grassberger, P., and I. Procaccia, "Estimation of the Kolmogorov Entropy from a Chaotic Signal," *Physical Review A*, Vol. 28, pp. 2591-2593.

Györfi, Lázló, Wolfgang Härdle, Pascal Sarda, and Phillipe Vieu, "Nonparametric Curve Estimation from Time Series," Berlin, New York, Tokyo: Springer-Verlag, Lecture Notes in Statistics, 1989.

Hsieh, David A., "Implications of Nonlinear Dynamics for Financial Risk Management," *Journal of Financial and Quantitative Analysis*, Vol. 28, No. 1, 1993, pp. 41-64.

Masry, Elias and Dag Tjóstheim, "Nonparametric Estimation and Identification of Nonlinear Arch Time Series: Strong Convergence and Asymptotic Normality," *Econometric Theory*, Vol. 11, 1995, pp. 258-289.

Miura, Ryozo, and H. Kishino, "Pricing of Bonds and Their Derivatives with Multifactor Stochastic Interest Rates: A Note," *Nonlinear and Convex Analysis in Economic Theory*, Lecture Notes in Economics and Mathematical Systems No.419, pp. 215-229, Springer-Verlag, 1995.

Morgan Guaranty Trust Company, "RiskMetrics™ Technical Document," New York, NY: J. P. Morgan, October 26, 1994.

Müller, Hans-Georg, "Nonparametric Regression Analysis of Longitudinal Data," Berlin, New York, Tokyo: Springer-Verlag, Lecture Notes in Statistics, 1988.

Phillips, P.C.B., "Optimal Inference in Cointegrated Systems," *Econometrica*, Vol. 59, No. 2, March 1991, pp. 283-306.

Ruppert, D., and M.P. Wand, "Multivariate Locally Weighted Least Squares Regression," *The Annals of Statistics*, Vol. 22, No. 3, 1994, pp. 1346-1370.

Shibata, Ritei, and Ryozo Miura, "Decomposition of Japanese Yen Interest Rates Data Through Local Regression," *Financial Engineering and the Japanese Markets*, Vol. 4, 1997, pp. 125-146.

Statistical Science, "S-PLUS Reference," Statistical Science, 1993.

Watson, Mark W., "Vector Autoregressive and Cointegration," Handbook of Econometrics, Vol. IV, 1994, pp. 2844-2919, Edited by Robert F. Engle and Daniel L. McFadden; Series Editors, Kenneth J. Arrow and Michael D. Intriligator, Amsterdam, New York, Tokyo: Elsevier.