

On Financial Time Series Decompositions with Applications to Volatility

Kjell Doksum¹, Ryozi Miura², and Hiroaki Yamauchi³

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, German Marks, Swiss Francs, and British Pounds in the unit of U.S. dollars. Similarities and differences in the trends between these currencies as well as their volatilities are discussed.

Key words : Autoregressive fit, BDS statistic, Decomposition, Exchange rates, Financial time series, LOWESS, Volatility.

JEL classification : C10, C14

¹ Department of Statistics, University of California, Berkeley, California, 94720-3860, United States.

² Faculty of Commerce, Hitotsubashi University, Kunitachi-shi, Tokyo, Japan.

³ Student in Doctoral program at Graduate School of Commerce, Hitotsubashi University, Kunitachi-shi, Tokyo, Japan.

Table of Contents.

I. Introduction	1--3
II. Decomposition of Currency Exchange Rates	4--41
II-1. Two-Step Symmetric Smoothing	4--19
(i) Long Term Trend	4
(ii) Short Term Trend	8
(iii) The Residual Series	11
(iv) The Decompositions	13
(v) Auto-Regressive Analysis of Residuals	16
II-2. Two-Step Predictive Smoothing	20--34
(i) Methodology of Predictive Smoothing	20
(ii) Long and Short Term Trends and the Residuals in Predictive Decompositions	21
(iii) Autoregressive Analysis of the Predictive Irregular Series	26
(iv) Prediction	32
II-3. Implications for Risk Measurement	35--41
(i) Behaviors of the four currencies against US dollar in three levels	37
(ii) Calculating VaR (Value at Risk)	38
(iii) Back Testing	41
III. Testing dependence in time series	42--45
III-1. BDS Statistics	42
III-2. Autoregressive Decomposition of the Residuals $\hat{I}(t)$: Symmetric Decomposition	44
III-3. Predictive Decomposition	45
IV. Volatility Measures and Their Time Series Decompositions	46--53
A. Volatility Measures Based on Returns	46
B. RiskMetrics™	49
C. Volatility Based on the Logarithmic Range	51
Appendix. Asymptotic Approximations	54--56
A. The Symmetric Regression Smoothers	54
B. The Predictive Regression Smoothers	56
References	57--58

I. Introduction

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

Our initial decomposition of the time series $Y(t)$, $t = 1, \dots, T$, is of the form

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

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, in the whole of this paper, that $Y(t)$ is a random variable with population mean $\mu(t)$, $t = 1, \dots, T$, and define $L(t)$, the theoretical long term trend, as a weighted average of $\mu(t)$ over a long time interval such as a quarter or a year. After subtracting out the long term trend, the short term trends $S(t)$ can be detected after removing the long term trend by taking a weighted average over a short time span. That is, the theoretical local features of variability can be measured by a short term weighted average of $\mu(t) - L(t)$. The left over fine grain theoretical variability 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 of $Y(t)$ 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 averages 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 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 Japanese Yen (JY), German Marks (GM), Swiss Francs (SF), and British Pounds (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 (Jan 5,1992) before

applying the decomposition. 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 German Marks and Swiss Francs exchange rates. We conjecture that the similarity between the movements of GM and SF is related to the macro-economic relations with the US 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 show 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 US dollar.

The chapters give further findings obtained by applying our methodology.

Figure 0 : Indexed Exchange Rates : JY, GM, SF, and BP

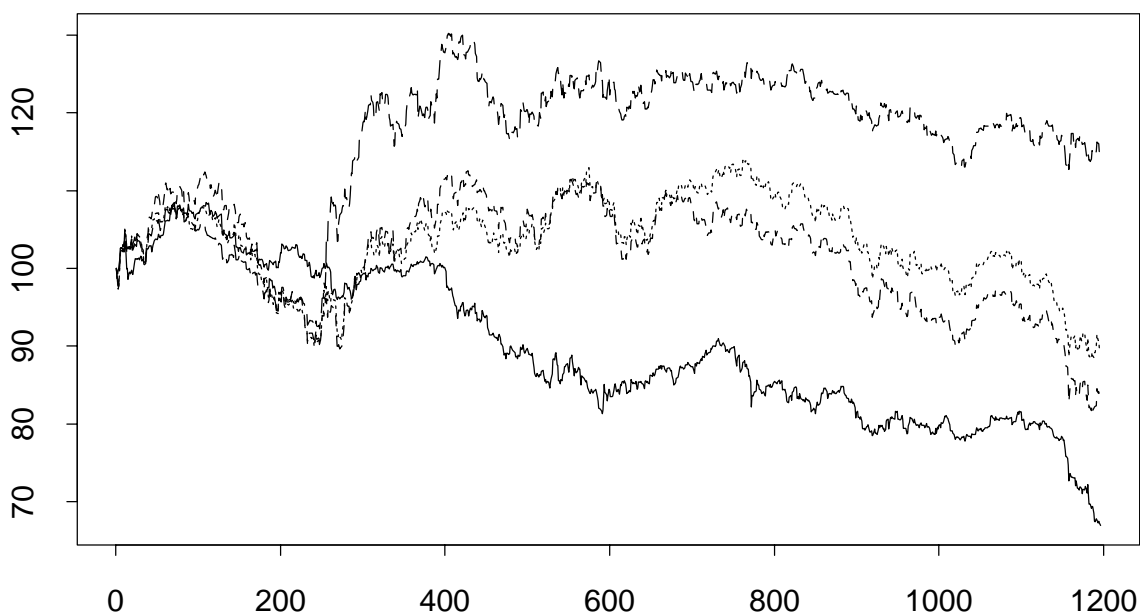


Figure 0. Indexed currency exchange rates of the Japanese Yen —, German Marks ·····, Swiss Francs - - - -, and British Pounds — — for the time period Jan 5, 1992 to April 14, 1995.

We consider two types of decompositions in Chapter II. 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 the data of 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 1 and 2 of Chapter II 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 3, Chapter II, we discuss the implications of our decomposition results for risk measurements and derivative pricing based on VaR(Value at Risk).

In order to check whether our decompositions have been successful in transforming the original time series into interpretable terms plus independent, identically distributed (i.i.d.) residuals, we need hypothesis-testing of the i.i.d. hypothesis. Here we use the Brock, Dechert, and Scheinkman (BDS)(1987) test statistic which is based on the correlation dimension in chaos theory. (This BDS statistics is also applied in Section 3 of Chapter II.)

In Chapter IV, we give two applications of the predictive decomposition to the assessment of weekday volatility in the exchange rates. Our first approach is based on introducing $V(t) = [X(t) - m(t)]^2$, where $X(t) = \text{Log}[Y(t)/Y(t-1)]$ 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 German Marks and Swiss Francs are close, while the Japanese Yen follows a more independent path. For the Japanese series we compare our method to J.P.Morgan's RiskMetrics™ and argue that our method gives a better representation of volatility. Our second approach to volatility is based on the logarithm of the days high to the days low, that is

$$W(t) = \text{Log} \frac{\text{High}}{\text{Low}} .$$

This measure has been shown to have some advantages over the squared deviations $V(t)$ as a volatility measure. We show how the predictive long and short term trends of $W(t)$ can be used to estimate conditional volatility, and again find strong similarities between German Marks and Swiss Francs, while the Japanese Yen is more individualistic.

In the appendix 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 locally linear predictive method introduced in Chapter II-2 tend to represent the "true" trend better than kernel method such as J.P.Morgan's RiskMetrics™ whenever the methods are applied to series where trends are present.

Chapter II. Decomposition of Currency Exchange Rates

Section II-1. Decomposition of Time Series in terms of Two-Step Symmetric Regression Smoothers

(i) Long Term Trend

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

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

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

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

where the kernel $K(u)$ is a symmetric density on the interval $[-1, 1]$, $2M-1$ is the length of the time span where $w_k(t)$ is positive, and $C = \sum_{k=t-M}^{t+M} K\left(\frac{k-t}{M}\right)$. Typically $\frac{C}{M}$ is very close to one since it can be approximated by the integral $\int_{-1}^1 K(u) du$, which equals 1 by the density assumption.

One possibility for a long term trend that has been considered in the literature (e.g. Beveridge and Nelson(1981), Enders(1994)) is the linear trend $\mu(t) = \alpha + \beta t$. If this is the true formula for $\mu(t)$, our definition of $L(t)$ will yield $L(t) = \alpha + \beta t$ for t between $t-M$ and $t+M$. This result can be extended to all of t by using 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 (1), and let a, b_1, \dots, b_d be the values of $\alpha, \beta_1, \dots, \beta_d$ that minimize the local weighted sum of squares.

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

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

$$L(t) = a = \sum_{k=t-M}^{t+M} w_k^*(t) \mu(k) . \quad (3)$$

where $w_k^*(t)$ are constants. In the case of a locally linear fit, where $d=1$,

$$w_k^*(t) = \frac{a_k(t)}{\sum_{k=t-M}^{t+M} a_k(t)}, \text{ with } a_k(t) = w_k(t) \sum_{i=t-M}^{t+M} w_i(t)(t-i)(k-i). \quad (4)$$

Locally polynomial smoothers of $\mu(t)$ have many desirable properties that have been demonstrated by Cleveland(1978), Cleveland and Devlin(1988), Fan(1993), Fan and Gijbels(1995), Ruppert and Wand(1995), 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 $d = p + 1$ locally polynomial long term trend function $L(t)$ equals $\mu(t)$ for all $t = 1, 2, \dots, T$.

Another interesting set of weights $w_k(t)$ is given by the set of "exponential weights"

$$w_k(t) = \frac{\lambda^{|t-k|}}{\sum_{k=1}^T \lambda^{|t-k|}}, \quad 0 < \lambda < 1.$$

These weights are symmetric versions, standardized to sum to one, of the weights used by J.P.Morgan's (1994) RiskMetrics™, 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, p.105). A very simple and easily understood set of weights are the "uniform moving average weights" which equal $(2r + 1)^{-1}$ for k in the interval $[t - r, t + r]$, and equal 0 elsewhere. In this case $L(t)$ is the uniform moving average

$$L(t) = (2r + 1)^{-1} \sum_{k=t-r}^{t+r} \mu(t).$$

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 (4) with $d = 1$ and

$$w_k(t) = M^2 - (k - t)^2, \quad t - M \leq k \leq t + M \\ = 0, \text{ otherwise}$$

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 1 gives the estimated long term trends as represented by the 91 and 181 day LOWESS SMOOTHERS for the exchange rates of Japanese Yen, German Marks, Swiss Francs, and British Pounds in U.S. dollars for the time period Jan 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 Japanese Yen, German Marks, Swiss Francs, and British Pounds. 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 US dollar in the four currencies when compared to the value on Jan 5, 1992. The four exchange rate values were all recorded at the same instance, 10 p.m., GMT.

Figure 1-A : Long Term Trends : Symmetric (91 days) : JY, GM, SF, and BP

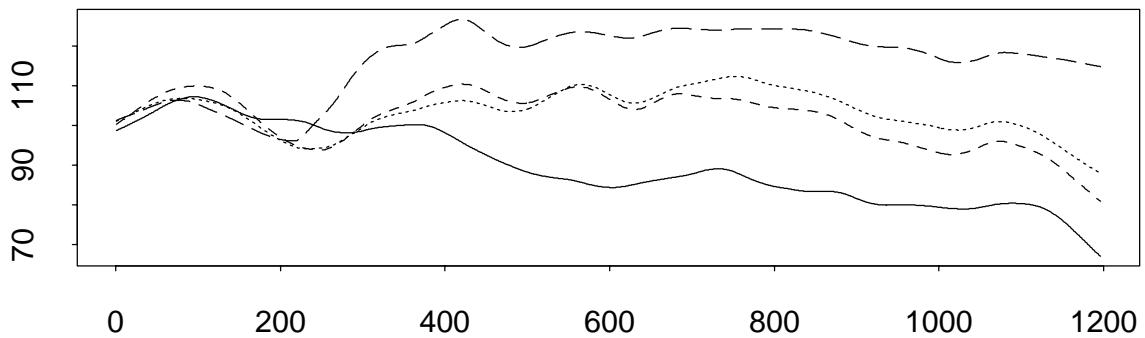


Figure 1-B : Long Term Trends : Symmetric (181 days) : JY, GM, SF, and BP

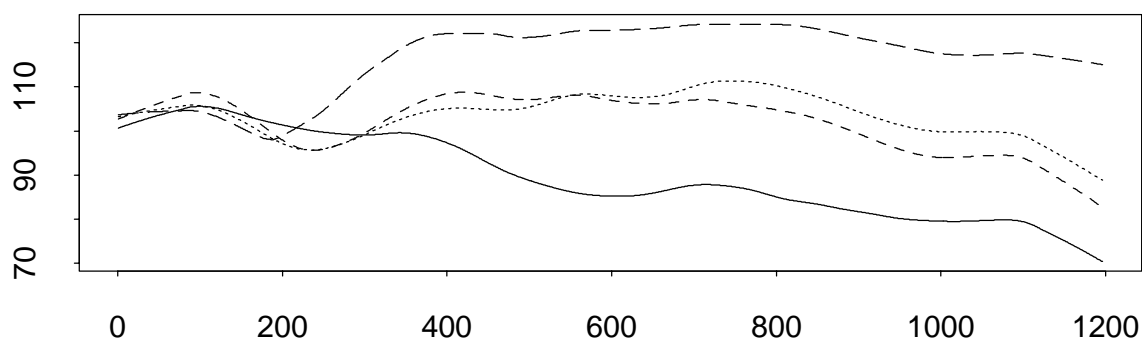


Figure 1. Long term trends of the Japanese Yen —, German Marks ·····, Swiss Francs - - - -, and British Pounds — — for the time period Jan 5, 1992 to April 14, 1995. Figure 1-A and 1-B give the 91 and 181 day LOWESS regression smoothers, respectively.

The graphs show that Swiss Francs and German Marks have very similar long term trends while the long term trends in Japanese Yen and British Pounds follow their own low and high paths, respectively. The “spike” in the British Pounds shortly after day 200 corresponds to the time when it was “cut loose” from other European currencies. Also note that starting near day 600, the path of the Japanese Yen is very similar to that of the German Marks and Swiss Francs. The 91 and 181 day long term trends are very similar except for the British Pounds shortly after day 200 when it was “cut loose” from other European currencies. The correlations are as follows.

Long Term Trend (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

< Table 1-1(a) >

Long Term Trend (181 days)	JY	GM	SF	BP
JY	1.00	0.09	0.53	-0.66
GM	0.09	1.00	0.85	0.51
SF	0.53	0.85	1.00	0.18
BP	-0.66	0.51	0.18	1.00

< Table 1-1(b) >

Daily Increments of Long Term Trend (91 days)	JY	GM	SF	BP
JY	1.00	0.56	0.61	0.29
GM	0.56	1.00	0.95	0.72
SF	0.61	0.95	1.00	0.73
BP	0.29	0.72	0.73	1.00

< Table 1-2(a) >

Daily Increments of Long Term Trend (181 days)	JY	GM	SF	BP
JY	1.00	0.53	0.59	0.26
GM	0.53	1.00	0.94	0.66
SF	0.59	0.94	1.00	0.66
BP	0.26	0.66	0.66	1.00

< Table 1-2(b) >

Table 1. Correlations between long term trends in the four currencies. Table 1-1(a) and Table 1-1(b) give respectively the correlations for LOWESS 91 day and 181 day long term trends. Table 1-2(a) and

Table1-2(b) give respectively the correlations for the daily increments of LOWESS 91 day and 181 day long term trends.

The correlation table gives a concise summary of the trends shown in the graphs. It also provides one surprise: The correlation 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. Note that the correlation of $dL(t)$ is very close to one for SF and GM.

(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

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

where the weights $v_k(t)$ are nonzero over a short time span. $S(t)$ depicts fluctuations which have been smoothed out in calculating $L(t)$ with the large 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 weeks), the seven day uniform moving average is without the day of the week effect.

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

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

Figure 2-A : Short Term Trends : Symmetric (91 days---7 days) : GM and SF

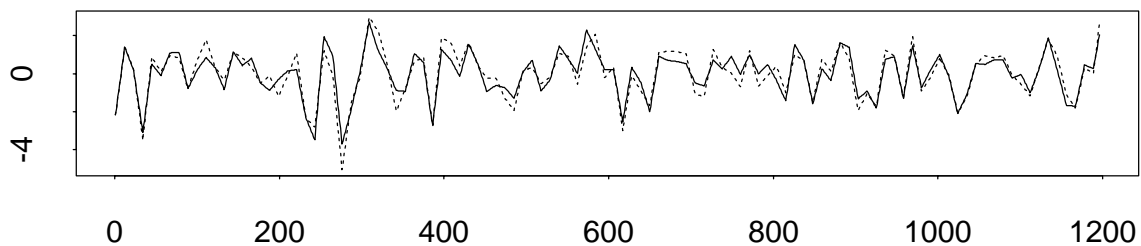


Figure 2-B : Short Term Trends : Symmetric (181 days---15 days) : GM and SF

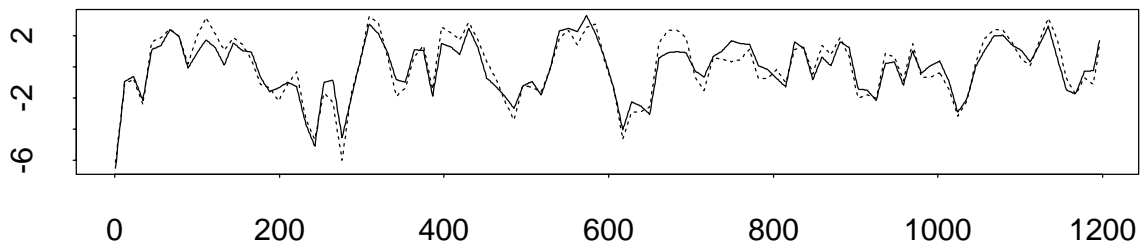


Figure 2-C : Short Term Trends : Symmetric (91 days---7 days) : JY and BP

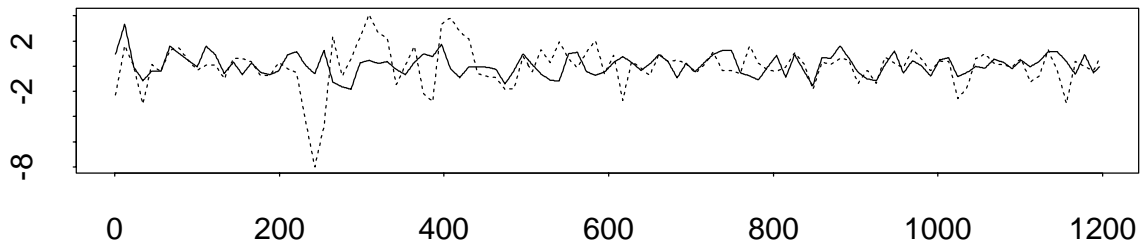


Figure 2-D : Short Term Trends : Symmetric (181 days---15 days) : JY and BP

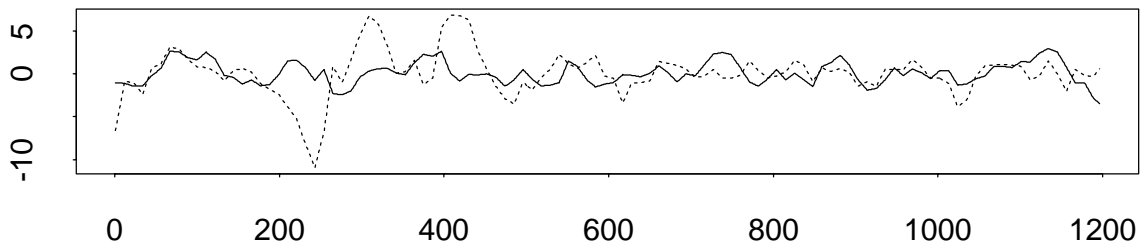


Figure 2. 2-A and 2-B give the short term trends of German Marks — and Swiss Francs using respectively 7 and 15 day short term trends based respectively on 91 and 181 day long term trends. 2-C and 2-D similarly give short term trends of Japanese Yen — and British Pounds

Figure 2 gives the estimated short term trends $\hat{S}(t)$ for the four currencies we consider. We see that the German Marks and Swiss Francs have very similar short term trends while the Japanese Yen and British Pounds do not. The correlations are as follows.

Short Term Trend (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

< Table 2-1(a) >

Short Term Trend (181---15days)	JY	GM	SF	BP
JY	1.00	0.42	0.46	0.15
GM	0.42	1.00	0.95	0.71
SF	0.46	0.95	1.00	0.70
BP	0.15	0.71	0.70	1.00

< Table 2-1(b) >

Daily Increments of Short Term Trend (91 --- 7 days)	JY	GM	SF	BP
JY	1.00	0.55	0.52	0.28
GM	0.55	1.00	0.93	0.69
SF	0.52	0.93	1.00	0.69
BP	0.28	0.69	0.69	1.00

< Table 2-2(a) >

Daily Increments of Short Term Trend (181---15days)	JY	GM	SF	BP
JY	1.00	0.41	0.41	0.11
GM	0.41	1.00	0.94	0.73
SF	0.41	0.94	1.00	0.70
BP	0.11	0.73	0.70	1.00

< Table 2-2(b) >

Table 2. Correlations between short term trends in the four currencies. (a) uses a 7 day short term trend based on a 91 days long term trend while (b) uses a 15 days short term trend based on a 181 day long term trend. Table 2-1 shows the correlations for the level of short term trends and Table 2-2 shows the correlations for their daily increments.

The relatively small correlations of Japanese Yen to the other currencies is an indication that it's short term movements are more influenced by national economic developments than the other currencies. In contrast to the long term trends, JY is now similarly related to SF and GM. Moreover, a strong local relationship between SF and GM is evident in the correlations tables.

(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

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

This series has an interpretation as a residual of a residual series, that is a re-residual. To see this rewrite $\hat{I}(t)$ as

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

We introduce $\hat{R}(t) = Y(t) - \hat{L}(t)$ as the empirical residuals after a long term fit. Then, by definition, $\hat{S}(t)$ is the nonparametric regression fit to $\hat{R}(t)$, and $\hat{R}(t) - \hat{S}(t)$ is the residual of the fit to the residual - a re-residual.

Figure 3 gives the empirical residual series of the four currencies. The same story as in Section II-1(i) and II-1(ii) emerges: The German Marks and Swiss Francs are closely correlated while the Japanese Yen and British Pounds show less of a relationship to the other currencies. The British Pound still shows a spike after it was cut loose from other European currencies. The correlations are as follows.

Residual (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

< Table 3(a) >

Residual (181---15days)	JY	GM	SF	BP
JY	1.00	0.50	0.48	0.38
GM	0.50	1.00	0.92	0.75
SF	0.48	0.92	1.00	0.69
BP	0.38	0.75	0.69	1.00

< Table 3(b) >

Table 3. Correlations between residuals of the four currencies. (a) uses a 7 day short term trend based on a 91 days long term trend while (b) uses a 15 days short term trend based on a 181 day long term trend.

Figure 3-A : Residuals : Symmetric (91 days---7 days) : GM and SF

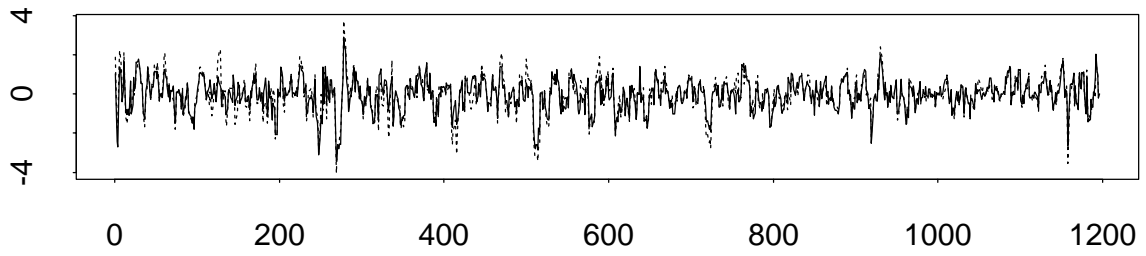


Figure 3-B : Residuals : Symmetric (181 days---15 days) : GM and SF

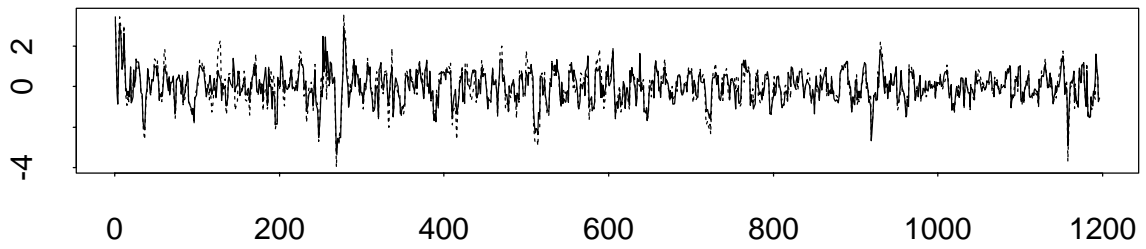


Figure 3-C : Residuals : Symmetric (91 days---7 days) : JY and BP

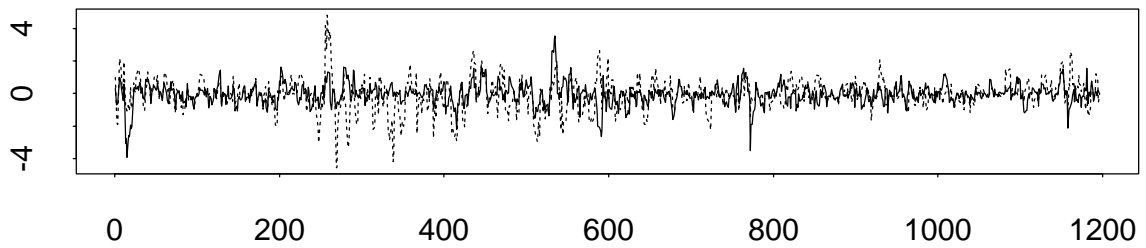


Figure 3-D : Residuals : Symmetric (181 days---15 days) : JY and BP

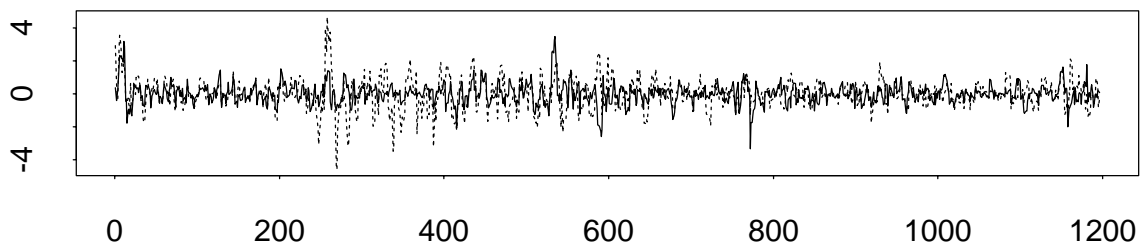


Figure 3. 3-A and 3-B give the residual series for German Marks — and Swiss Francs using (A) 91 and 7 days long and short term trends and (B) 181 and 15 days long and short term trends. 3-C and

3-D similarly give residual series for the Japanese Yen ——— and British Pounds - - - - -.

(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 residual series is (nearly) "pure noise", that is whether it consists of (nearly) independent identically distributed random variables. We will return to this question in Section IV.

Another 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 measure the extent to which each series provide "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 4.

Correlation for (91,7) days 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) days 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	0.99	-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) days 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) days 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

< Table 4 >

Table 4. Correlations between the terms in the decompositions of four exchange rate series.

The correlations between $\hat{L}(t)$ and $\hat{S}(t)$ in Table 4 show that JY is decomposed better than the other currencies, since for JY this correlation is close to zero. For the time spans 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, in Figure 4, we compare 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 vector autoregressive fit to $\hat{I}(t)$ for the (91,7) symmetric decomposition. 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 $I(t)$.

Figure 4-A : Box Plots of Increments of Exchange Rates : $Y(t) - Y(t-1)$: Scaled

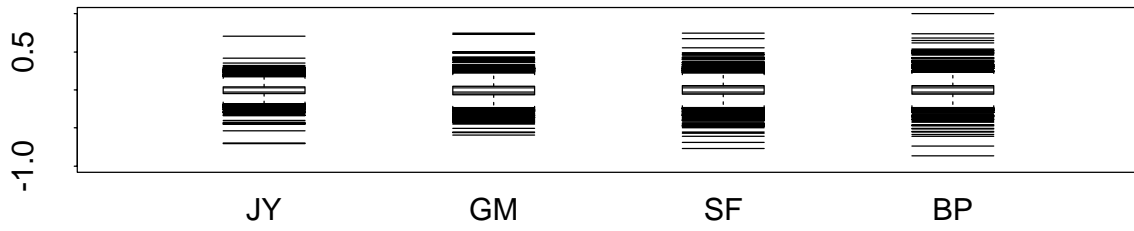


Figure 4-B : Box Plots of Increments of Long Term Trends (91 days) : $\hat{L}(t) - \hat{L}(t-1)$: Symmetric : Scaled

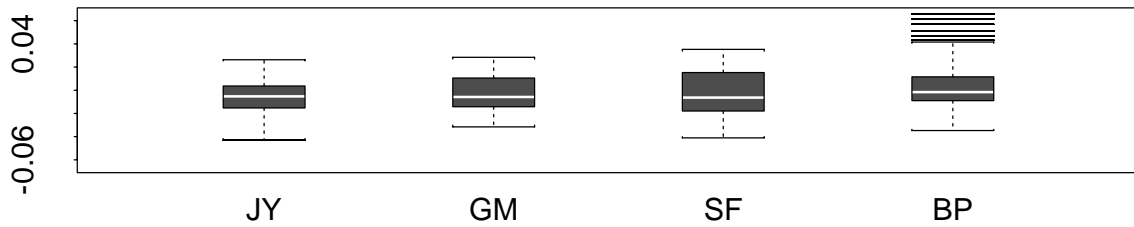


Figure 4-C : Box Plots of Increments of Short Term Trends (91 days---7 days) : $\hat{S}(t) - \hat{S}(t-1)$: Symmetric : Scaled

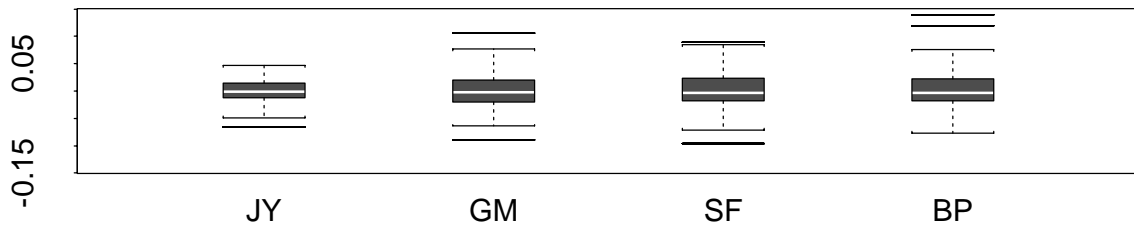


Figure 4-D : Box Plots of Irregular Term (91 days---7 days) : $\hat{I}(t)$: Symmetric : Scaled

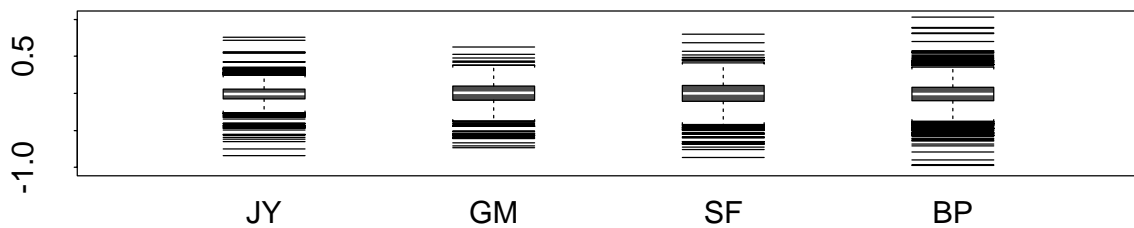


Figure 4-E : Box Plots of Increments of Irregular Term (91 days---7 days) : $\hat{I}(t) - \hat{I}(t - 1)$:
Symmetric : Scaled

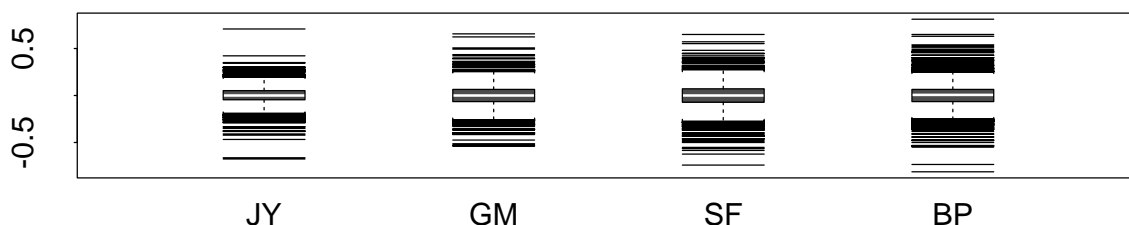


Figure 4. Box plots of daily changes of (A) the exchange rates $Y(t)$, (B) the long term trends $\hat{L}(t)$, (C) the short term trends $\hat{S}(t)$, (D) the irregular series $\hat{I}(t)$, and (E) the increments of the irregular series using 91 and 7 days long and short term trends.

Remark. 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 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 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) Auto-Regressive 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, which will be defined later in Section III-1 of Chapter III, to the residuals and the results are shown in Table 17 in that later section. The results there suggest that properties of the residuals are very much like that of an independent and identically distributed sequence of

variables.

The coefficient matrices below suggest an univariate autoregressive fit of order one, for Yen. This leads to the fitted model

$$\hat{I}(t) = 0.724\hat{I}(t-1) + \varepsilon_1(t)$$

with prediction error variance 0.200 as compared to 0.198 for the MAR(4) model fit. The BDS statistics for this $\varepsilon_1(t)$ will be shown in Section III-2. This suggests that a good fit to the Yen series is the parsimonious model

$$\begin{aligned} Y(t) &= L(t) + S(t) + 0.724I(t-1) + \varepsilon_1(t) \\ &= 0.724Y(t-1) + [L(t) - 0.724L(t-1)] + [S(t) - 0.724S(t-1)] + \varepsilon_1(t). \end{aligned}$$

A similar reasoning applied to German Marks yield the model

$$Y(t) = 0.704Y(t-1) + [L(t) - 0.704L(t-1)] + [S(t) - 0.704S(t-1)] + \varepsilon_1(t)$$

with prediction error variance 0.313.

The model of four-variate autoregressive fit is below.

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

$$\begin{aligned} \text{, where } \mathbf{I}(t) &= \begin{bmatrix} \mathbf{I}(t) \text{ of JY} \\ \mathbf{I}(t) \text{ of GM} \\ \mathbf{I}(t) \text{ of SF} \\ \mathbf{I}(t) \text{ of BP} \end{bmatrix}, \quad \mathbf{e}(t) = \begin{bmatrix} \varepsilon(t) \text{ of JY} \\ \varepsilon(t) \text{ of GM} \\ \varepsilon(t) \text{ of SF} \\ \varepsilon(t) \text{ of BP} \end{bmatrix}, \quad \mathbf{C}_1 = \begin{bmatrix} 0.793 & -0.076 & 0.000 & -0.001 \\ -0.035 & 0.590 & 0.098 & 0.047 \\ -0.044 & 0.014 & 0.755 & 0.013 \\ -0.022 & -0.224 & 0.063 & 0.856 \end{bmatrix}, \\ \mathbf{C}_2 &= \begin{bmatrix} -0.041 & 0.029 & 0.000 & 0.026 \\ -0.001 & 0.085 & -0.079 & 0.030 \\ 0.010 & -0.022 & -0.079 & 0.072 \\ -0.020 & 0.024 & -0.016 & 0.066 \end{bmatrix}, \quad \mathbf{C}_3 = \begin{bmatrix} 0.003 & 0.081 & -0.058 & -0.037 \\ 0.091 & 0.113 & -0.163 & -0.044 \\ 0.046 & 0.099 & -0.099 & -0.051 \\ 0.075 & 0.302 & -0.269 & -0.116 \end{bmatrix}, \text{ and} \\ \mathbf{C}_4 &= \begin{bmatrix} -0.038 & -0.089 & 0.081 & -0.002 \\ -0.043 & -0.248 & 0.220 & 0.023 \\ 0.008 & -0.196 & 0.142 & 0.016 \\ -0.019 & -0.305 & 0.282 & -0.014 \end{bmatrix}. \end{aligned}$$

Prediction Error Variance	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

AIC (subtracted minimum number of AIC)	order 0	order 1	order 2	order 3	order 4	order 5
	3636.348	7.967	2.523	5.980	0.000	3.055
	order 6	order 7	order 8	order 9	order 10	order 11
	11.842	6.607	6.256	31.930	25.875	28.330

< Table 5-1 >

Table 5-1. Prediction error variances, and AIC for the order 4 autoregressive fit to the residuals $\hat{I}(t)$ of the (91,7) symmetric decomposition.

		Prin. Component 1	Prin. Component 2	Prin. Component 3	Prin. Component 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
	BP	0.495	-0.362	0.781	0.117

< Table 5-2 >

Table 5-2. Principal components analysis for the residuals $\varepsilon_4(t)$ of the order 4 vector autoregressive fit to the residuals $\hat{I}(t)$ of the (91,7) symmetric decomposition.

Thus, 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 fitting using principal component analysis. The results are shown in Table 5-2. We see that while all the four 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 common factors mostly except the fourth principal component whose explanatory power is however very small (as small as 2.5%).

Figure 5 : Box Plots of Residuals of MAR fit to Irregular Term (91 days---7 days) : $\hat{I}(t)$:
Symmetric : Scaled

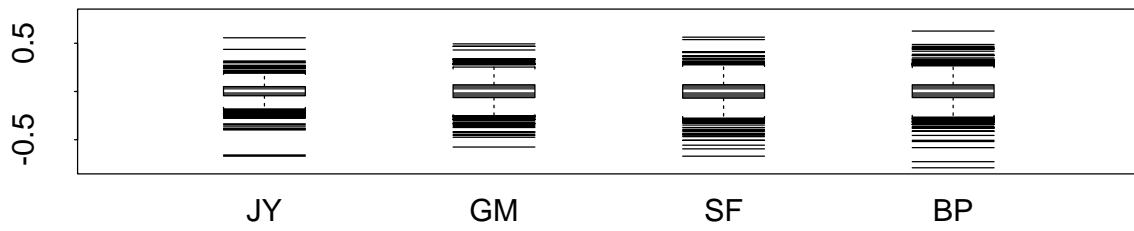


Figure 5. Box plot of residuals of autoregressive fit to the irregular series of symmetric decomposition.

Section II-2. Two-Step Decomposition of Time Series with Predictive Smoothing

(i) Methodology of Predictive Smoothing

The methods in the preceding section provide a 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 . Thus we want to measure the long and short term trends as well as the residuals at time t using measurements up to the 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

$$L_p(t) = \sum_{k=t-TL}^t w_k(t) \mu(k), S_p(t) = \sum_{k=t-TS}^t v_k(t) [\mu(k) - L_p(k)],$$

$$\text{and } I_p(t) = Y(t) - [L_p(t) + S_p(t)]$$
(5)

where $w_k(t)$ and $v_k(t)$ are as in Section II-1. For instance we might use

$$w_k(t) = \frac{\lambda_1^{t-k}}{\sum_{k=1}^t \lambda_1^{t-k}}, v_k(t) = \frac{\lambda_2^{t-k}}{\sum_{k=1}^t \lambda_2^{t-k}}, k = 1, \dots, t$$

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 (1994) RiskMetrics™ 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 (5) are obtained by replacing $\mu(k)$ by $Y(k)$ in the formulae (5).

In the case that only lagged values $Y(t-k)$, $k \geq 0$, 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-1 is superior. See the Appendix for the mathematical expression of this bias.

Figure 6-A : Long Term Trends (Predictive and Symmetric : 91 days) : JY

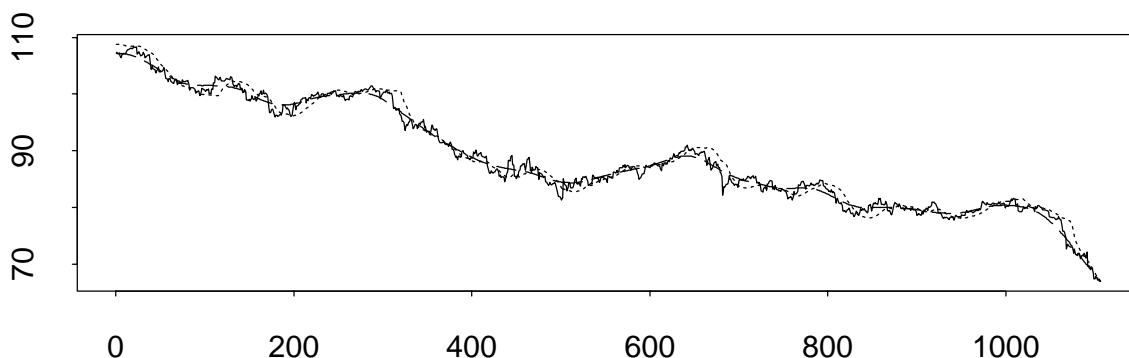


Figure 6-B : Long Term Trends (Predictive and Symmetric : 91 days) : GM

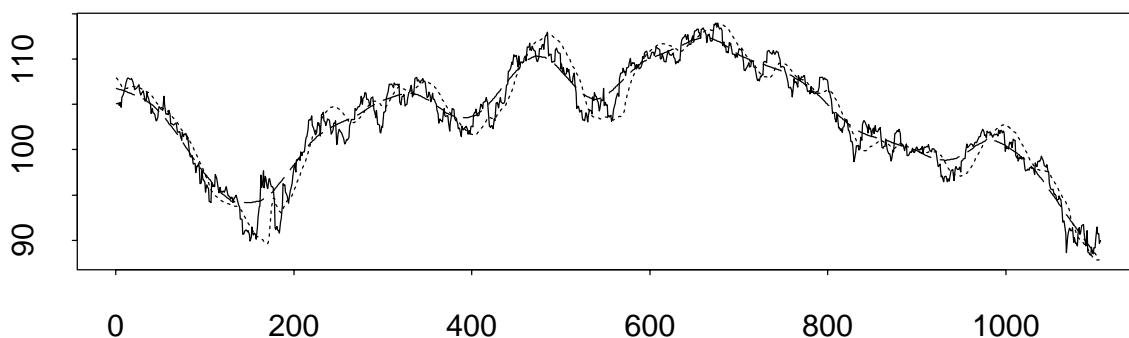


Figure 6-A. Indexed currency exchange rate of Japanese Yen ——— and the symmetric — — and predictive ······ long term trends series.

Figure 6-B. Indexed currency exchange rate of German Marks ——— and the symmetric — — and predictive ······ long term trends series.

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

Figure 6 compares the past (ex-post, symmetric) and predictive long term trends $\hat{L}(t)$ and $\hat{L}_p(t)$, based on the locally linear method (weights), for the Japanese Yen and German Marks 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) overshoots and undershoots the “true” trend $\hat{L}(t)$ much more than the locally linear $\hat{L}_p(t)$.

We further examined the correlation of the levels and increments of $L_p(t)$. Table 6-1 shows the correlation coefficients of $L_p(t)$ of the four currencies and Table 6-2 shows that of lag-one (daily) increments. We see that correlation of JY and BP is very small (almost uncorrelated) and

correlation of JY and GM is around 0.4 which is somewhat smaller than that of JY and SF. The correlation of GM and SF is very high, as high as 0.85. These results are quite concordant with those based on symmetric smoothing noted in Section II-1-(i).

Long Term Trend (91 days)	JY	GM	SF	BP
JY	1.00	0.07	0.46	-0.46
GM	0.07	1.00	0.87	0.64
SF	0.46	0.87	1.00	0.41
BP	-0.46	0.64	0.41	1.00

< Table 6-1(a) >

Long Term Trend (181 days)	JY	GM	SF	BP
JY	1.00	-0.05	0.31	-0.30
GM	-0.05	1.00	0.87	0.78
SF	0.31	0.87	1.00	0.66
BP	-0.30	0.78	0.66	1.00

< Table 6-1(b) >

Daily Increments of Long Term Trend (91 days)	JY	GM	SF	BP
JY	1.00	0.33	0.41	0.09
GM	0.33	1.00	0.85	0.37
SF	0.41	0.85	1.00	0.43
BP	0.09	0.37	0.43	1.00

< Table 6-2(a) >

Daily Increments of Long Term Trend (181 days)	JY	GM	SF	BP
JY	1.00	0.39	0.43	0.13
GM	0.53	1.00	0.90	0.64
SF	0.59	0.94	1.00	0.49
BP	0.26	0.66	0.66	1.00

< Table 6-2(b) >

Table 6-1 Correlations of long term predictive trends for the four currencies.

Table 6-2 Correlations of lag-1 increments of long term predictive trends for the four currencies.

Figure 7 and 8 give 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 (Jan 5, 1995) of the exchange rate series while the short term and residual series start 91+7=98 days after Jan 5, 1992. All the predictive series stop on April 14, 1995, one day before the last value in the series. The results are very similar to the symmetric decompositions of Section II-1 except the long and short term trends in the predictive decompositions are more volatile.

We can also show that the predictive decompositions are made up of nearly orthogonal terms. This is seen from Table 9 which shows correlation coefficients of the predictively decomposed components. Before that, we check the correlations of levels and increments of $S_p(t)$ of the four currencies, which are shown in Table 7-1 and 7-2. Again the correlations are quite similar to that of the symmetric decompositions, but interestingly the differences of JY's correlations with GM(or SF) and BP are much less. It was 0.55-0.28 in the symmetric case, but it is 0.39-0.23 in this predictive case. This means that it is more difficult in the predictive case to distinguish between GM(or SF) and BP than in the symmetric case. The same effects are also seen in $I_p(t)$. Table 8 shows correlation coefficients of $I_p(t)$ of the four currencies, where we see that the correlations

are smaller (by 0.1) than those in Table 3 in the previous section. These are the effects of not using the data posterior to the time t . The choice of the time span such as (91,7) and (181,15) also affects the resulting magnitude of the correlations. We will see later in this section that time span (91,7) is superior to (181,15) in terms of prediction errors.

Figure 7-A : Short Term Trends : Predictive (91 days---7 days) : GM and SF

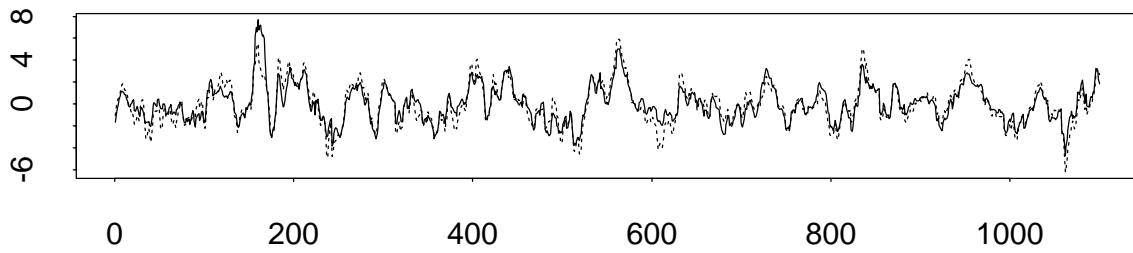


Figure 7-B : Short Term Trends : Predictive (91 days---7 days) : JY and BP

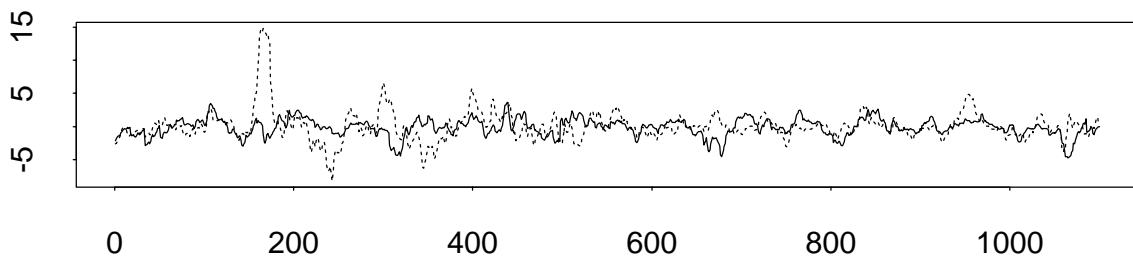


Figure 7-C : Residuals : Predictive (91 days---7 days) : GM and SF

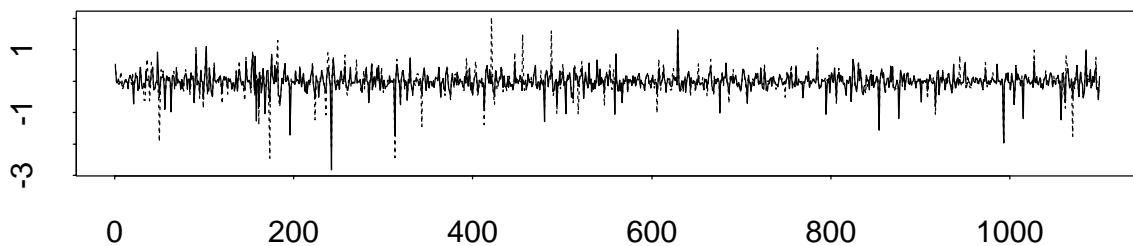


Figure 7-D : Residuals : Predictive (91 days---7 days) : JY and BP

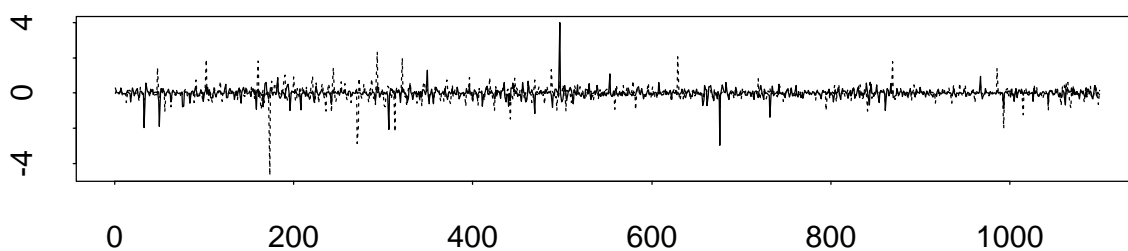


Figure 7. 7-A gives the short term trends of German Marks ——— and Swiss Francs using the 91 long term and 7 day short term predictive decompositions. 7-B similarly gives short term trends for Japanese Yen ——— and British Pounds 7-C and 7-D give the corresponding residuals.

Figure 8 : Long Term Trends : Predictive (91 days) : JY, GM, SF,and BP

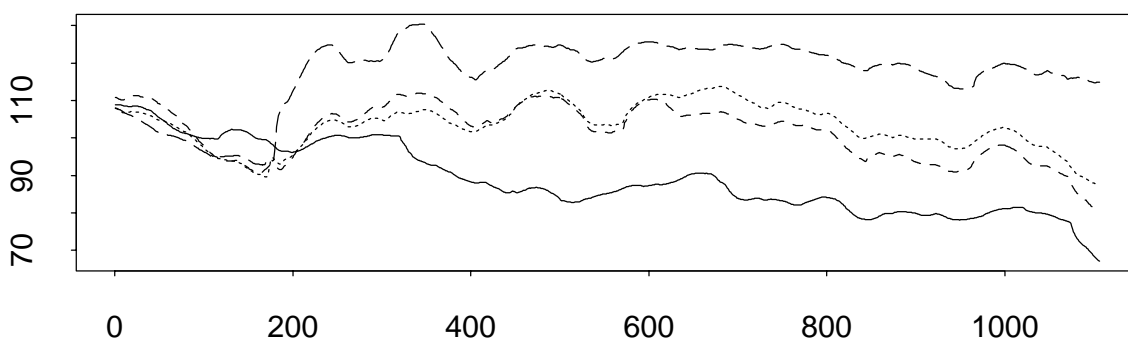


Figure 8. Predictive long term trends of the Japanese Yen ———, German Marks , Swiss Francs ---- ,and British Pounds — — for the time period April 4, 1992 to April 13, 1995. The graphs are 91 day LOWESS regression smoothers.

Short Term Trend (91--- 7 days)	JY	GM	SF	BP
JY	1.00	0.41	0.43	0.11
GM	0.41	1.00	0.91	0.66
SF	0.43	0.91	1.00	0.60
BP	0.11	0.66	0.60	1.00

< Table 7-1(a) >

Short Term Trend (181---15days)	JY	GM	SF	BP
JY	1.00	0.40	0.42	0.08
GM	0.40	1.00	0.94	0.74
SF	0.42	0.94	1.00	0.69
BP	0.08	0.74	0.69	1.00

< Table 7-1(b) >

Daily Increments of Short Term Trend (91 --- 7 days)	JY	GM	SF	BP
JY	1.00	0.39	0.40	0.23
GM	0.39	1.00	0.82	0.60
SF	0.40	0.82	1.00	0.56
BP	0.23	0.60	0.56	1.00

< Table 7-2(a) >

Daily Increments of Short Term Trend (181 --- 5days)	JY	GM	SF	BP
JY	1.00	0.34	0.32	0.22
GM	0.34	1.00	0.85	0.63
SF	0.32	0.85	1.00	0.59
BP	0.22	0.63	0.59	1.00

< Table 7-2(b) >

Table 7-1 Correlations of short term predictive trends for the four currencies.

Table 7-2 Correlations of lag-1 increments of short term predictive trends for the four currencies.

Residual (91 --- 7 days)	JY	GM	SF	BP
JY	1.00	0.37	0.37	0.27
GM	0.37	1.00	0.80	0.61
SF	0.37	0.80	1.00	0.59
BP	0.27	0.61	0.59	1.00

< Table 8(a) >

Residual (181---15days)	JY	GM	SF	BP
JY	1.00	0.41	0.37	0.31
GM	0.41	1.00	0.89	0.70
SF	0.37	0.89	1.00	0.66
BP	0.31	0.70	0.66	1.00

< Table 8(b) >

Table 8 Correlations of residuals for the four currencies.

The following Table 9 shows that the predictive decomposition is very successful in achieving orthogonality for JY and is also quite successful for other three European currencies except the remaining small correlations between $L_p(t)$ and $S_p(t)$.

Correlation for (91,7) days 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) days 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) days 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) days 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

< Table 9 >

Table 9 Correlations of components in the predictive decompositions.

(iii) Autoregressive Analysis of the Predictive Irregular Series

Table 10-1 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. Table 10-1 suggests an univariate order 2 autoregressive fit to the Japanese Yen exchange rate series:

$$Y(t) = \hat{L}_p(t) + \hat{S}_p(t) + 0.137\hat{I}_p(t-1) - 0.197\hat{I}_p(t-2) + \hat{\varepsilon}_p(t) \quad (8)$$

where the residual variance is 0.076 which should be compared with the variance 0.080 of $\hat{I}_p(t)$. This shows that the decomposition $Y(t) = \hat{L}_p(t) + \hat{S}_p(t) + \text{i.i.d. noise}$ is very successful in that the reduction in variance by going to the more complex model (8) is only $0.080 - 0.076 = 0.004$.

The four variate autoregressive fitted model is:

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

$$\text{, where } \mathbf{I}_p(t) = \begin{bmatrix} \mathbf{I}(t) \text{ of JY} \\ \mathbf{I}(t) \text{ of GM} \\ \mathbf{I}(t) \text{ of SF} \\ \mathbf{I}(t) \text{ of BP} \end{bmatrix}, \mathbf{e}_p(t) = \begin{bmatrix} \varepsilon(t) \text{ of JY} \\ \varepsilon(t) \text{ of GM} \\ \varepsilon(t) \text{ of SF} \\ \varepsilon(t) \text{ of BP} \end{bmatrix}, \mathbf{C}_1 = \begin{bmatrix} 0.149 & -0.061 & 0.008 & -0.046 \\ -0.024 & -0.010 & 0.055 & 0.007 \\ -0.034 & -0.036 & 0.176 & -0.036 \\ -0.014 & -0.051 & -0.078 & 0.188 \end{bmatrix},$$

$$\mathbf{C}_2 = \begin{bmatrix} -0.183 & -0.069 & 0.041 & 0.020 \\ -0.008 & -0.110 & -0.056 & -0.014 \\ -0.000 & -0.075 & -0.173 & 0.029 \\ -0.022 & -0.154 & 0.075 & -0.116 \end{bmatrix}, \text{ and } \mathbf{C}_3 = \begin{bmatrix} -0.110 & 0.066 & -0.098 & -0.011 \\ 0.012 & -0.043 & -0.150 & 0.004 \\ -0.024 & 0.058 & -0.238 & 0.028 \\ 0.016 & 0.047 & -0.142 & -0.078 \end{bmatrix}.$$

Prediction Error Variance	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

AIC (subtracted minimum number of AIC)	order 0	order 1	order 2	order 3	order 4	order 5
	241.895	175.135	53.566	0.000	16.668	19.834

< Table 10-1 >

Table 10-1. Prediction error variances, and AIC for the order 3 autoregressive fit to the residuals $\hat{I}_p(t)$ of the (91,7) predictive decomposition.

		Prin. Component 1	Prin. Component 2	Prin. Component 3	Prin. Component 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
	BP	0.497	-0.317	0.808	0.010

< Table 10-2 >

Table 10-2. Principal components analysis for the residuals of the order 3 vector autoregressive fit to the residuals $\hat{I}_p(t)$ of the (91,7) predictive decomposition.

Remark 1. The (91,7) predictive decomposition uses 91 days prior to t to make the long term prediction $\hat{L}_p(t)$ and seven days prior to t to make the short term prediction $\hat{S}_p(t)$. This short term predictor can be regarded as a "local autoregressive" fit. It is a weighted linear fit using seven lagged values. Autoregressive techniques applied to the residuals $\hat{I}_p(t)$ are teasing out possible overall (nonlocal) linear dependencies on lagged values. For instance, with Japanese Yen there were some slight autoregressive dependencies of $\hat{I}_p(t)$ on the 2 first lagged values.

Figure 9-A : Box Plots of Increments of Long Term Trends (91 days) : $\hat{L}_p(t) - \hat{L}_p(t-1)$:
 Predictive : Scaled

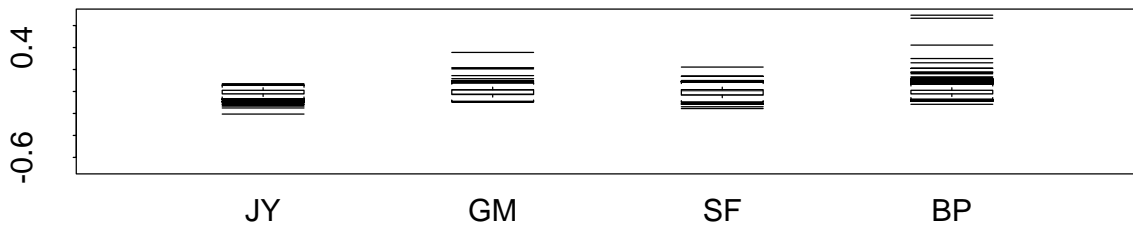


Figure 9-B : Box Plots of Increments of Short Term Trends (91 days---7 days) : $\hat{S}_p(t) - \hat{S}_p(t-1)$:
 Predictive : Scaled

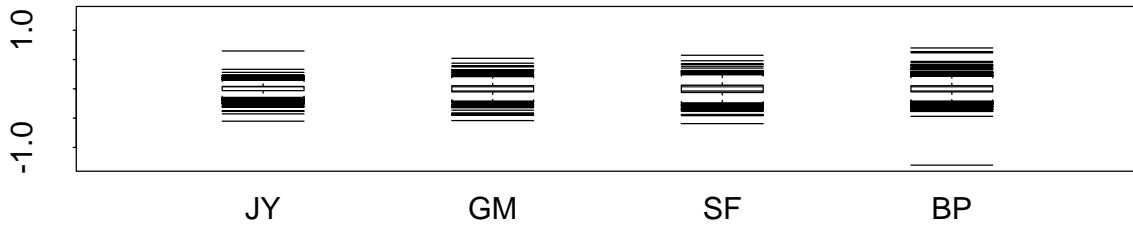


Figure 9-C : Box Plots of Irregular Term (91 days---7 days) : $\hat{I}_p(t)$: Predictive : Scaled

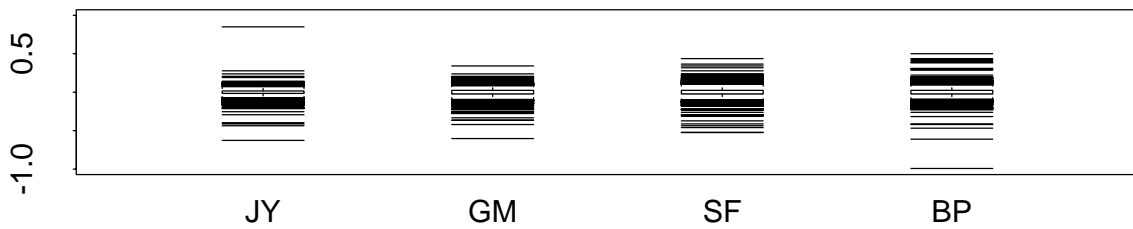


Figure 9-D : Box Plots of Increments of Irregular Term (91 days---7 days) : $\hat{I}_p(t) - \hat{I}_p(t-1)$:
 Predictive : Scaled

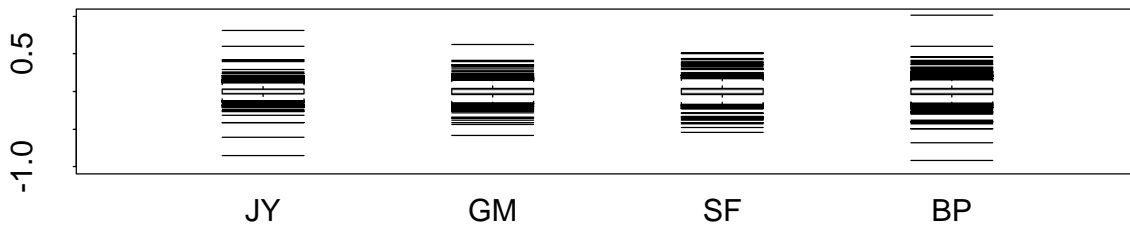


Figure 9-E : Box Plots of Residuals of MAR fit to Irregular Term (91 days---7 days) : $\hat{I}_p(t)$:
 Predictive : Scaled

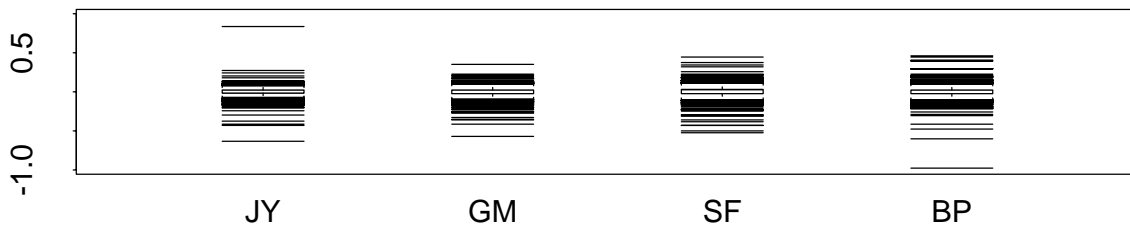


Figure 9. Box plots of (A) daily increments of the long term trends $\hat{L}_p(t)$, (B) the increments of the short term trends $\hat{S}_p(t)$, (C) the irregular series $\hat{I}(t)$, (D) the increments of the irregular series, and (E) the autoregressive residuals using 91 and 7 days long and short term trends.

Next we consider the (181,15) predictive decomposition where $L_p(t)$ depends on 181 lagged values and $S_p(t)$ depends on 15 lagged values. In this case we find a different result from the (91,7) decomposition. The multivariate autoregressive analysis shown in Table 11-1 yields an order 7 (Akaike) model. The fitted model is below.

$$\mathbf{I}_p(t) = \mathbf{C}_1 \cdot \mathbf{I}_p(t-1) + \mathbf{C}_2 \cdot \mathbf{I}_p(t-2) + \dots + \mathbf{C}_7 \cdot \mathbf{I}_p(t-7) + \mathbf{e}_p(t)$$

$$\text{, where } \mathbf{I}_p(t) = \begin{bmatrix} \mathbf{I}(t) \text{ of JY} \\ \mathbf{I}(t) \text{ of GM} \\ \mathbf{I}(t) \text{ of SF} \\ \mathbf{I}(t) \text{ of BP} \end{bmatrix}, \mathbf{e}_p(t) = \begin{bmatrix} \boldsymbol{\varepsilon}(t) \text{ of JY} \\ \boldsymbol{\varepsilon}(t) \text{ of GM} \\ \boldsymbol{\varepsilon}(t) \text{ of SF} \\ \boldsymbol{\varepsilon}(t) \text{ of BP} \end{bmatrix}, \mathbf{C}_1 = \begin{bmatrix} 0.573 & -0.020 & -0.037 & 0.004 \\ -0.060 & 0.527 & 0.027 & 0.016 \\ -0.051 & 0.141 & 0.491 & -0.019 \\ -0.027 & -0.024 & -0.058 & 0.598 \end{bmatrix},$$

$$\mathbf{C}_2 = \begin{bmatrix} -0.100 & 0.008 & 0.033 & 0.025 \\ 0.016 & -0.050 & -0.008 & 0.064 \\ 0.018 & -0.061 & -0.075 & 0.070 \\ -0.049 & -0.010 & 0.012 & 0.069 \end{bmatrix}, \mathbf{C}_3 = \begin{bmatrix} -0.025 & 0.112 & -0.081 & -0.056 \\ 0.043 & 0.099 & -0.129 & -0.061 \\ 0.012 & 0.093 & -0.109 & -0.043 \\ 0.063 & 0.442 & -0.328 & -0.195 \end{bmatrix},$$

$$\mathbf{C}_4 = \begin{bmatrix} -0.052 & -0.057 & 0.074 & 0.004 \\ -0.047 & -0.205 & 0.176 & 0.034 \\ 0.022 & -0.247 & 0.205 & -0.008 \\ -0.010 & -0.321 & 0.237 & 0.001 \end{bmatrix}, \quad \mathbf{C}_5 = \begin{bmatrix} -0.079 & -0.025 & 0.007 & 0.008 \\ 0.007 & -0.058 & -0.028 & 0.016 \\ 0.018 & 0.093 & -0.177 & 0.035 \\ 0.000 & 0.068 & -0.084 & -0.004 \end{bmatrix}, \\
\mathbf{C}_6 = \begin{bmatrix} -0.012 & 0.015 & 0.033 & -0.074 \\ 0.055 & -0.039 & -0.049 & 0.002 \\ 0.015 & -0.018 & -0.062 & 0.012 \\ 0.092 & 0.020 & -0.005 & -0.041 \end{bmatrix}, \quad \text{and} \quad \mathbf{C}_7 = \begin{bmatrix} -0.146 & 0.013 & -0.044 & 0.024 \\ -0.013 & -0.169 & 0.061 & 0.005 \\ 0.049 & -0.106 & -0.015 & -0.014 \\ 0.006 & -0.077 & -0.050 & -0.086 \end{bmatrix}.$$

Prediction Error Variance	JY	GM	SF	BP
JY	0.175	0.102	0.104	0.091
GM	0.102	0.257	0.245	0.216
SF	0.104	0.245	0.307	0.224
BP	0.091	0.216	0.224	0.375

AIC (subtracted minimum number of AIC)	order 0	order 1	order 2	order 3	order 4	order 5
	1689.859	223.242	192.898	148.232	121.648	65.295
	order 6	order 7	order 8	order 9	order 10	order 11
	46.055	0.000	15.197	24.033	30.451	28.094
order 12	order 13	order 14	order 15	order 16	order 17	
43.807	48.014	59.391	65.189	77.799	87.109	
order 18	order 19	order 20	order 21	order 22	order 23	
103.639	119.188	134.328	147.678	169.488	187.582	

< Table 11-1 >

Table 11-1. Prediction error variances, and AIC for the order 7 autoregressive fit to the residuals $\hat{\mathbf{I}}_p(t)$ of the (181,15) predictive decomposition.

		Prin. Component 1	Prin. Component 2	Prin. Component 3	Prin. Component 4
EigenValue		2.806	0.698	0.372	0.125
Contribution		70.14%	17.44%	9.29%	3.13%
Cumulative Contribution		70.14%	87.58%	96.87%	100.00%
Eigenvectors	JY	0.382	0.917	0.111	0.036
	GM	0.557	-0.161	-0.343	-0.739
	SF	0.546	-0.198	-0.463	0.669
	BP	0.496	-0.307	0.810	0.065

< Table 11-2 >

Table 11-2. Principal components analysis for the residuals of the order 7 vector autoregressive fit to the residuals $\hat{\mathbf{I}}_p(t)$ of the (181,15) predictive decomposition.

Remark 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 Tjøstheim (1995).

Remark 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 did change a little. The Akaike Information Criteria selects an order 6 multivariate autoregressive fit and the lag one through three coefficients explains most of the variability with one interesting exception: The lag 6 (weekly) coefficient for British Pounds is large for German Marks and Swiss Francs. The multivariate lag one coefficient for Yen is 0.149 when weekends are included and 0.125 when weekends are excluded. The fitted model is:

$$\mathbf{I}_P(t) = \mathbf{C}_1 \cdot \mathbf{I}_P(t-1) + \mathbf{C}_2 \cdot \mathbf{I}_P(t-2) + \dots + \mathbf{C}_7 \cdot \mathbf{I}_P(t-7) + \mathbf{e}_P(t)$$

$$\text{, where } \mathbf{I}_P(t) = \begin{bmatrix} \mathbf{I}(t) \text{ of JY} \\ \mathbf{I}(t) \text{ of GM} \\ \mathbf{I}(t) \text{ of SF} \\ \mathbf{I}(t) \text{ of BP} \end{bmatrix}, \mathbf{e}_P(t) = \begin{bmatrix} \varepsilon(t) \text{ of JY} \\ \varepsilon(t) \text{ of GM} \\ \varepsilon(t) \text{ of SF} \\ \varepsilon(t) \text{ of BP} \end{bmatrix}, \mathbf{C}_1 = \begin{bmatrix} 0.125 & -0.019 & -0.016 & -0.057 \\ -0.038 & 0.092 & 0.029 & -0.035 \\ -0.040 & 0.002 & 0.163 & -0.099 \\ -0.032 & 0.006 & -0.101 & 0.149 \end{bmatrix},$$

$$\mathbf{C}_2 = \begin{bmatrix} -0.198 & 0.092 & -0.097 & 0.012 \\ 0.000 & -0.149 & -0.052 & -0.017 \\ 0.001 & -0.071 & -0.190 & -0.003 \\ 0.022 & -0.021 & -0.023 & -0.189 \end{bmatrix}, \mathbf{C}_3 = \begin{bmatrix} -0.134 & 0.088 & -0.102 & -0.038 \\ -0.001 & -0.164 & -0.063 & -0.011 \\ 0.011 & 0.042 & -0.254 & -0.023 \\ 0.019 & -0.018 & -0.091 & -0.105 \end{bmatrix},$$

$$\mathbf{C}_4 = \begin{bmatrix} -0.075 & -0.106 & 0.115 & -0.028 \\ 0.022 & -0.026 & -0.030 & -0.025 \\ 0.003 & -0.019 & -0.005 & -0.064 \\ 0.032 & 0.022 & 0.031 & -0.097 \end{bmatrix}, \mathbf{C}_5 = \begin{bmatrix} -0.088 & 0.146 & -0.145 & -0.005 \\ -0.050 & -0.107 & -0.027 & 0.046 \\ 0.036 & -0.024 & -0.141 & 0.034 \\ -0.038 & -0.019 & 0.052 & -0.045 \end{bmatrix}, \text{and}$$

$$\mathbf{C}_6 = \begin{bmatrix} -0.090 & -0.003 & -0.058 & 0.002 \\ 0.005 & 0.021 & -0.075 & -0.086 \\ 0.013 & -0.042 & -0.031 & -0.072 \\ -0.001 & 0.288 & -0.294 & -0.156 \end{bmatrix}.$$

Prediction Error Variance	JY	GM	SF	BP
JY	0.000	0.000	0.000	0.000
GM	0.000	0.000	0.000	0.000
SF	0.000	0.000	0.000	0.000
BP	0.000	0.000	0.000	0.000

AIC (subtracted minimum number of AIC)	order 0	order 1	order 2	order 3	order 4	order 5
	231.285	172.404	90.077	43.439	49.600	22.589
	order 6	order 7	order 8	order 9	order 10	order 11
	0.000	4.475	2.979	17.940	32.831	47.102
order 12	order 13	order 14	order 15	order 16	order 17	
69.445	81.815	90.372	90.526	101.688	126.703	
order 18	order 19	order 20	order 21	order 22		
137.719	136.961	141.391	156.097	178.150		

< Table 12 >

Table 12. Prediction error variances, and AIC for the order 7 autoregressive fit to the residuals $\hat{\mathbf{I}}_P(t)$ of the (181,15) predictive decomposition.

(iv) Prediction

We now consider the problem of predicting $Y(t+1)$ from past values $Y(k)$, $k \leq t$. Suppose we use $\hat{L}_p(t)$ initially to predict $Y(t+1)$. The unknown prediction error is then $Y(t+1) - \hat{L}_p(t)$. This error can be assessed by using a weighted average of known lagged prediction errors, that is by

$$S_p^*(t) = \sum_{k=t-TS}^{t-1} v_k(t) [Y(k+1) - \hat{L}_p(k)].$$

Now $\hat{L}_p(t) + S_p^*(t)$ is the revised predicted value of $Y(t+1)$. If our decomposition has been successful, then the residual time series

$$\hat{R}_p(t+1) = Y(t+1) - [\hat{L}_p(t) + S_p^*(t)]$$

is nearly stationary and can be further decomposed using autoregressive techniques. That is, we can express $\hat{R}_p(t+1)$ as

$$\hat{R}_p(t+1) = \sum_{j=-k}^0 \hat{C}_j \hat{R}_p(t-j) + \text{error}.$$

, where $\hat{R}_p(t) = \begin{bmatrix} \hat{R}_p(t) \text{ of JY} \\ \hat{R}_p(t) \text{ of GM} \\ \hat{R}_p(t) \text{ of SF} \\ \hat{R}_p(t) \text{ of BP} \end{bmatrix}$, $\hat{C}_1 = \begin{bmatrix} 0.469 & -0.102 & -0.013 & 0.013 \\ -0.041 & 0.237 & 0.097 & 0.020 \\ -0.049 & -0.008 & 0.406 & 0.040 \\ -0.051 & -0.216 & 0.045 & 0.506 \end{bmatrix}$,

$$\hat{C}_2 = \begin{bmatrix} -0.188 & 0.064 & 0.004 & -0.009 \\ -0.014 & 0.023 & -0.123 & 0.032 \\ -0.019 & 0.066 & -0.208 & 0.006 \\ -0.052 & 0.029 & -0.016 & -0.030 \end{bmatrix}, \text{ and } \hat{C}_3 = \begin{bmatrix} -0.084 & 0.117 & -0.080 & -0.040 \\ 0.052 & 0.087 & -0.141 & -0.042 \\ 0.027 & 0.181 & -0.212 & -0.029 \\ 0.023 & 0.241 & -0.220 & -0.131 \end{bmatrix}.$$

The estimated coefficients $\{\hat{C}_j\}$ of the fit are based on $\{\hat{R}_p(t) ; t = 99,100, \dots, 1196\}$. This yields the final prediction

$$\hat{Y}(1197) = \hat{L}_p(1196) + S_p^*(1196) + \sum_{j=-k}^0 \hat{C}_j \hat{R}_p(1196-j)$$

for $Y(1197)$. Table 13 gives the covariance matrix of the error term,

$Y(t+1) - \left[\hat{L}_p(t) + S_p^*(t) + \sum_{j=-k}^0 \hat{C}_j \hat{R}_p(t-j) \right]$, and variances of predicted values of $Y(t+1)$, that is,

$$\hat{L}_p(t) + S_p^*(t) + \sum_{j=-k}^0 \hat{C}_j \hat{R}_p(t-j).$$

Variance	JY	GM	SF	BP
JY	0.263	0.160	0.177	0.133
GM	0.160	0.412	0.397	0.334
SF	0.177	0.397	0.509	0.335
BP	0.133	0.334	0.335	0.538

< Table 13-1 >

Variance	JY	GM	SF	BP
JY	82.432	3.718	29.824	-37.319
GM	3.718	34.491	35.979	32.953
SF	29.824	35.979	50.324	25.655
BP	-37.319	32.953	25.655	77.426

< Table 13-2 >

Table 13. Variance-covariances Matrices of (1) $Y(t+1) - \left[\hat{L}_p(t) + S_p^*(t) + \sum_{j=-k}^0 \hat{C}_j \hat{R}_p(t-j) \right]$
and (2) $\hat{L}_p(t) + S_p^*(t) + \sum_{j=-k}^0 \hat{C}_j \hat{R}_p(t-j)$.

Note from the coefficient matrices that for Japanese Yen all the coefficients are small except for the lag1 and 2 Japanese Yen exchange rate coefficients. This suggests a very parsimonious order 2 univariate autoregressive model for Japanese Yen ($t = 99, 100, \dots, 1196$),

$$\hat{R}_p(t+1) = 0.425\hat{R}_p(t) - 0.189\hat{R}_p(t-1) + \text{error}$$

with error term mean -0.0003 and variance 0.271. This compares with an error term mean -0.0005 and variance 0.263 for the order 3 multivariate autoregressive fit. The prediction at time point 1196 based on the univariate AR(2) decomposition of $\hat{R}_p(t+1)$ for Japanese Yen is thus

$$\hat{Y}(1197) = \hat{L}_p(1196) + S_p^*(1196) + 0.425\hat{R}_p(1196) - 0.189\hat{R}_p(1195).$$

We also computed the prediction for time points prior to 1197 starting with $t=108$ and an autoregressive fit to $\hat{R}_p(t)$ based on the 9 values prior to 108. Next we predicted $Y(109)$ using 10 values prior to 109 for the autoregressive fit, and so on up to $t=1197$. The resulting error means and variances are given in Table 14-1. Note that the \hat{C}_j coefficients in the autoregressive fits are recomputed each time using the expanding windows $99, 100, \dots, t$ and thus depend on t . They are too numerous to be given here.

	JY	GM	SF	BP
Mean	-0.0047	-0.0022	0.0074	-0.0042

Variance	JY	GM	SF	BP
JY	0.288	0.169	0.194	0.145
GM	0.169	0.453	0.436	0.360
SF	0.194	0.436	0.564	0.359
BP	0.145	0.360	0.359	0.585

< Table 14-1 >

Table 14-1. Means and covariance matrix of error term using expanding windows to

compute the coefficients $\hat{\mathbf{C}}_{j,t}$ in the multivariate autoregressive fit.

We also used fifty day windows to compute the coefficients, $\hat{\mathbf{C}}_{j,t}$. That is, to produce our predictor of $Y(t+1)$, we used the autoregressive fit based on the fifty days prior to $t+1$. The resulting predictor

$$\hat{Y}(t+1) = \hat{L}_p(t) + S_p^*(t) + \sum_{j=-k}^0 \hat{\mathbf{C}}_{j,t} \hat{\mathbf{R}}_p(t-j)$$

for Japanese Yen has error mean 0.0013 and variance 0.324. The error means and variances of the predictor based on the fifty day windows for the four currencies are given in Table 14-2.

	JY	GM	SF	BP
Mean	0.0013	0.0007	0.0077	0.0023

Variance	JY	GM	SF	BP
JY	0.324	0.181	0.195	0.154
GM	0.181	0.488	0.482	0.387
SF	0.195	0.482	0.622	0.402
BP	0.154	0.387	0.402	0.648

< Table 14-2 >

Table 14-2. Means and covariance matrix of error term using fifty day windows to compute the coefficients $\hat{\mathbf{C}}_{j,t}$ in the multivariate autoregressive fit.

Chapter II-3. Implications for Risk Measurement

Based on the decompositions done 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 relations 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 variance of the whole time series movements into nearly orthogonal parts. This interpretation for three components, i.e. three levels of uncertainties, is still 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. We have not done this yet. So the terms, macro, micro and daily moves used in the following suggestions for applications are based on our heuristic interpretations.

Table 15-(a) and 15-(b) shows the matrix of correlation coefficients of daily increment of the decomposed time-series components, $L(t)$ and $S(t)$, and the final AR-residuals, for the four currencies. They are shown for both of symmetric and predictive smoothing cases, respectively.

	dL(t)	dL(t)	dL(t)	dL(t)	dS(t)	dS(t)	dS(t)	dS(t)	ARres	ARres	ARres	ARres
	JY	GM	SF	BP	JY	GM	SF	BP	JY	GM	SF	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:BP	-0.01	-0.06	-0.05	-0.05	0.01	0.02	0.01	0.05	0.42	0.76	0.71	1.00

Table 15-(a). Correlations between $dL(t)$, $dS(t)$ and AR-residuals of the Symmetric decomposition (91,7)

	dLp(t) JY	dLp(t) GM	dLp(t) SF	dLp(t) BP	dSp(t) JY	dSp(t) GM	dSp(t) SF	dSp(t) BP	ARres JY	ARres GM	ARres SF	ARres BP
dLp(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
dLp(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
dLp(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
dLp(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
dSp(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
dSp(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
dSp(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
dSp(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:BP	0.03	-0.01	-0.01	-0.06	0.21	0.39	0.33	0.36	0.27	0.60	0.60	1.00

Table 15-(b). Correlations between $dL_p(t)$, $dS_p(t)$, and AR-residuals of the Predictive decomposition (91,7)

From the Box-Plots for daily increments of $Y(t)$, $L(t)$, $S(t)$, and $I(t)$, and for $I(t)$ itself and the final residuals, (i.e., Figure 4), we see 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 that 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, -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 0.2, respectively in our predictive (91-7) decomposition. (These numbers are based on rescaled unit used for Figure 4, 5, and 9). For other currencies see Table 16-(a) and 16-(b). where 10% and 90% percentiles of $dL(t)$, $dS(t)$, and AR-residual for each currency are shown. These figures are important because these relative-magnitude counts on average as relative contributions for the whole increment of $Y(t)$ that is "Risk".

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.

dL(t) JY		dL(t) GM		dL(t) SF		dL(t) BP	
10%	90%	10%	90%	10%	90%	10%	90%
-0.1005	0.0495	-0.1216	0.0982	-0.1410	0.1116	-0.0885	0.1197
dS(t) JY		dS(t) GM		dS(t) SF		dS(t) BP	
10%	90%	10%	90%	10%	90%	10%	90%
-0.1211	0.1179	-0.1733	0.2061	-0.1978	0.2231	-0.2353	0.1986
AR-residuals : JY		AR-residuals : GM		AR-residuals : SF		AR-residuals : BP	
10%	90%	10%	90%	10%	90%	10%	90%
-0.4954	0.4481	-0.6578	0.6294	-0.6991	0.6917	-0.7181	0.6684

Table 16-(a). 10% and 90% percentiles of $dL(t)$, $dS(t)$, and AR-residuals of four currencies based on the Symmetric decomposition (91,7)

dL _p (t) JY		dL _p (t) GM		dL _p (t) SF		dL _p (t) BP	
10%	90%	10%	90%	10%	90%	10%	90%
-0.1864	0.0891	-0.1843	0.1713	-0.2229	0.2044	-0.1683	0.1954
dS _p (t) JY		dS _p (t) GM		dS _p (t) SF		dS _p (t) BP	
10%	90%	10%	90%	10%	90%	10%	90%
-0.3862	0.4147	-0.5306	0.5542	-0.6062	0.6040	-0.6060	0.5912
AR-residuals : JY		AR-residuals : GM		AR-residuals : SF		AR-residuals : BP	
10%	90%	10%	90%	10%	90%	10%	90%
-0.1935	0.2017	-0.2380	0.2666	-0.2816	0.2959	-0.2943	0.2878

Table 16-(b). 10% and 90% percentiles of $dL_p(t)$, $dS_p(t)$, and AR-residuals of four currencies based on the Predictive decomposition (91,7)

(i) Behaviors of the four currencies against US dollar in three levels

Our currency exchange rates data represent the relative weakness of each currency against US dollars.

The correlation coefficients of $dL(t)$ (increment of $L(t)$, $L(t+1) - L(t)$) and $dS(t)$ ($= S(t+1) - S(t)$) of Japanese Yen 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 a lot smaller for BP. These numbers decrease in the predictive smoothing case. Note that $dL(t)$ of JY and BP are uncorrelated in the predictive smoothing case though it is correlated a little (0.29) in symmetric smoothing. More detailed correlations for final AR-residuals have been provided in the previous sections (II-1(v) and II-2(iii)) with principal component analysis.

It is possible under our decompositions to relate the component-wise correlations to

correlation of the whole $dY(t)$ for any two currencies using Table 15-(a) and 15-(b). We do not explicitly write it down here since it is straight forward.

The gaps in the numbers between the two smoothings reflect the lack of trading information in the predictive smoothing.

A more detailed study will require the analysis of the relationship with other data including macro- and micro-economic indices. We leave it for later research.

(ii) Calculating VaR (Value at Risk)

Assuming that the long term trend does not have any sudden changes, it should be possible to reduce the value of a standard deviation of $dY(t)$. Note that $dY(t) = dL_p(t) + dS_p(t) + dI_p(t)$ and our idea is the $dL_p(t)$ may be almost a predictable constant under the assumption.

At a time point t , our prediction $\hat{Y}(t+1)$ of $Y(t+1)$ 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, $dY(t)$ is decomposed as follows:

$$\begin{aligned} dY(t) &= Y(t+1) - Y(t) \\ &= \{Y(t+1) - \hat{Y}(t+1)\} + \{\hat{Y}(t+1) - Y(t)\} \\ &= dL_p(t) + dS_p(t) + (a_t - 1)I_p(t) + b_t I_p(t-1) + \dots \end{aligned}$$

Regarding $dL_p(t)$ as almost a constant, the variance of $dY(t)$ is almost that of sum of $dS_p(t)$ and I_p terms. (See Figure 10)

Following this intuition, we tried the measure VaR. Our estimate of 1% percentile, or Value at Risk, will be almost

$$\{\hat{Y}(t+1) - Y(t)\} - 2.33\sigma \quad (\sigma \text{ stands for standard deviation of } [dS_p(t) + I_p \text{ terms}]).$$

The assumed steady movements of long term trend of currency exchange rates will be realistic if the macro-economic and political relations of the two countries are steady. When steady, we do not need to take the variabilities caused by the unsteady relations.

In order to see how our VaR works, we calculated our VaR for 201 days in a row, and compared with two other standard VaR based on equally weighted sample variance and J.P.Morgan's exponentially weighted sample variance.

Our VaR was wild in a sense that the actual $dY(t)$ went below our VaR, 22 times within 201 days while this happened 3 times for J.P.Morgan's.

Our intuitive idea needs to be more polished so that this excess frequency be more compatible with 1% chance. Our effort is to have a small and efficient $\hat{\sigma}$ by squeezing out the practical uncertainties for VaR. Our $\hat{\sigma}$ is smaller than other two in most of the case. This is what we thought to happen.

Figure 10-A : $dL_p(t)$: JY

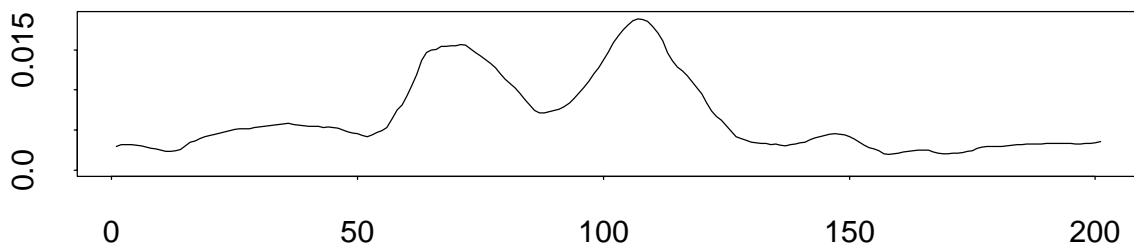


Figure 10-B : $dS_p(t)$: JY

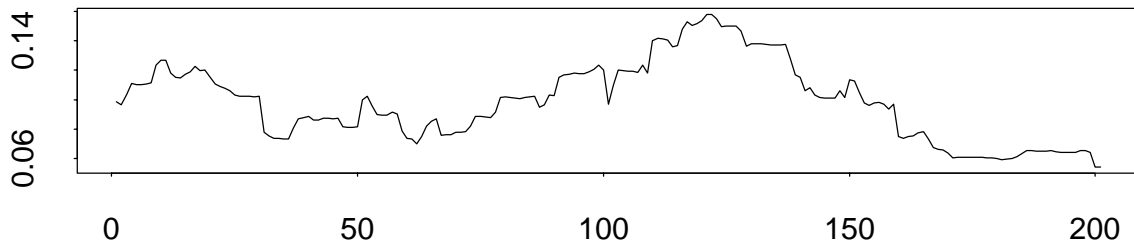


Figure 10-C : $(a_t - 1)I_p(t) + b_t I_p(t-1) + \dots$: JY

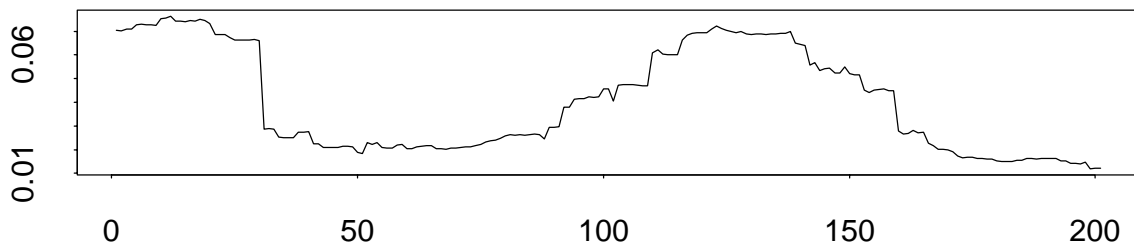


Figure 10. 10-A and 10-B give daily increments of 91 days long term trends and 91 and 7 days shot term trends for Japanese Yen, respectively. 10-C gives I_p term.

Figure 11-A : VaR : Equally Weighted : $\text{Log} \frac{Y(t+1)}{Y(t)} : JY$

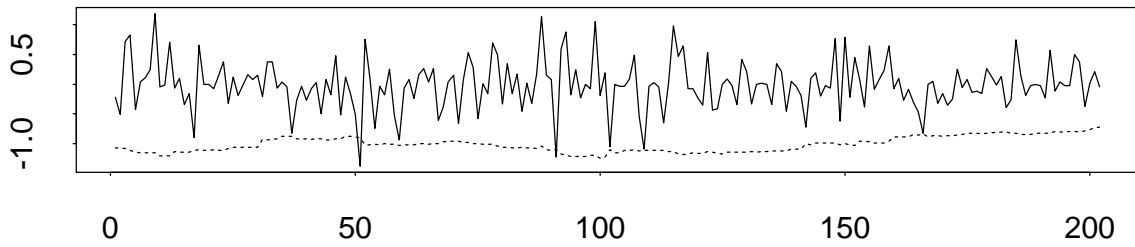


Figure 11-B : VaR : Exponentially Weighted ($\lambda = 0.94$) : $\text{Log} \frac{Y(t+1)}{Y(t)} : JY$

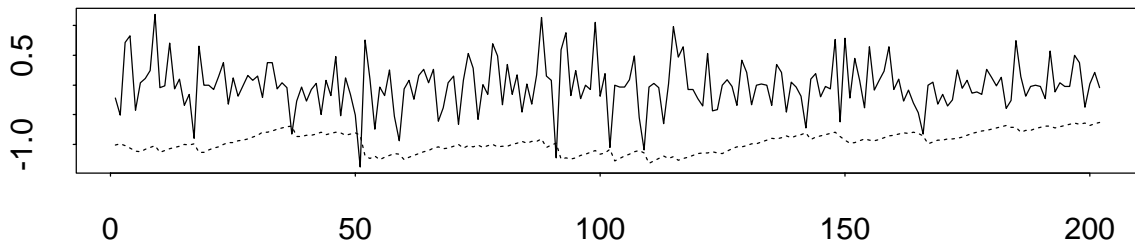


Figure 11-C : VaR : Exponentially Weighted ($\lambda = 0.94$) : $\hat{Y}(t+1) - Y(t) : JY$

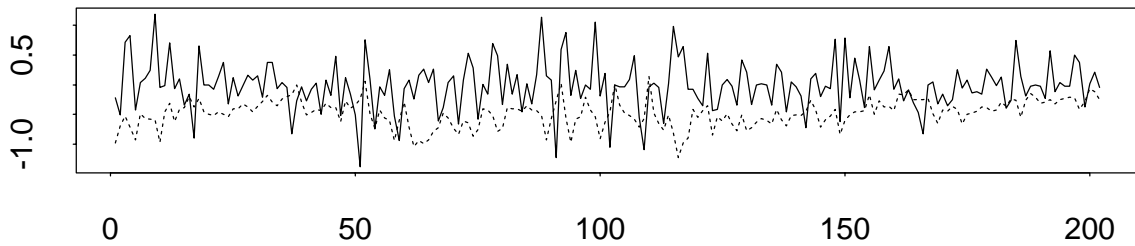


Figure 11. Plots of actual daily increments of $Y(t)$ — and VaR Figure 11-A and B are based on the differences of logarithm of $Y(t)$ using equally weights and exponentially weights, respectively. Figure 11-C is based on the differences of predicted value of tomorrow and today's value, that is , $\hat{Y}(t+1) - Y(t)$ using exponentially weights.

Figure 12 : Comparison of VaR : JY

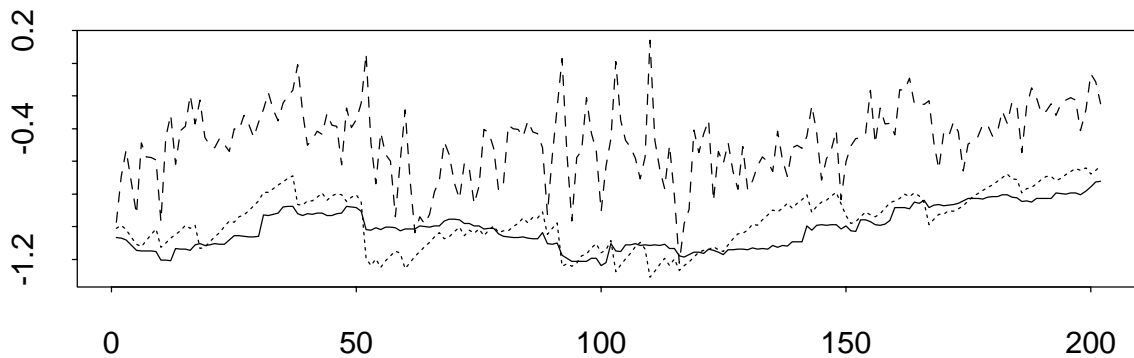


Figure 12. This figure gives the comparison of VaR which are based on differences of logarithm with equally weights ———, based on differences of logarithm with exponentially weights - - - - - , and based on $\hat{Y}(t+1) - Y(t)$ with exponentially weights - , respectively.

(iii) Back Testing

Back-Testing of 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 at the time. They count how many times the realized values went over the estimated 1% percentile, i.e., VaR.

Our two-step smoothing technique provides the decomposition of the variability from two viewpoints. So it will be interesting to see and compare the time series behavior of VaR with our decomposed $L(t)$, $S(t)$, and $I(t)$. And 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 decompositions), and we may also check what was known to the practitioners at the time in the predictive way.

The same remarks apply to the 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 smoothings and the implied volatilities provided among the traders in the markets

We leave these interesting problems for later research.

Chapter III. Testing and measuring dependence within and between time series

III-1. The BDS statistic

The graphs of the residual series in Figure 3, 7-C, and 7-D indicate a dependence of $\hat{I}(t)$ on its own lagged values. One way to measure such dependence in a general time series $X(t)$ is by the BDS (Brock, Dechert and Scheinkman (1987)) correlation dimension statistic

$$C_{m,T}(\delta) = \frac{2}{T_m(T_m - 1)} \sum_{t < s} I_\delta(X_t^m, X_s^m)$$

where $X_t^m = (X(t), \dots, X(t + m - 1))$, $T_m = t + (m - 1)$, and $I_\delta(X_t^m, X_s^m)$ is the indicator function of the event $\|X_t^m - X_s^m\| < \delta$. Here $\|\cdot\|$ is the supremum norm. An estimate of the standard deviation of $C_{m,T}(\delta)$ is

$$\sigma_{m,T}^2(\varepsilon) = 4 \left[K_T^m + 2 \sum_{j=1}^{m-1} K_T^{m-j} C^{2j} + (m-1)^2 C^{2m} - m^2 K_T C^{2m-2} \right]$$

$$\text{,where } K_T(\varepsilon) = \frac{6}{T_m(T_m - 1)(T_m - 2)} \sum_{t < s < r} \frac{\begin{bmatrix} I_\delta(X_t^m, X_s^m) I_\delta(X_s^m, X_r^m) \\ + I_\delta(X_t^m, X_r^m) I_\delta(X_r^m, X_s^m) \\ + I_\delta(X_s^m, X_t^m) I_\delta(X_t^m, X_r^m) \end{bmatrix}}{3} \text{ and } C = C_{1,T}(\delta).$$

It is known (Brock and Dechert(1988), Brock, Hsieh, and LeBaron (1991)) that under the null hypothesis of independent identically distributed $X(t)$,

$$\text{BDS} = \frac{\sqrt{T} (C_{m,T}(\hat{\delta}) - C_{1,T}^m(\hat{\delta}))}{\sigma_{m,T}(\hat{\delta})} \quad (6)$$

has asymptotically as $T \rightarrow \infty$ (and approximately for $T \geq 500$) a standard normal distribution. In this paper we will take $\hat{\delta} = k\hat{\sigma}$, where $\hat{\sigma}$ is the sample standard deviation of the series being analyzed, and $k = 0.50, 0.75, 1.00, 1.25, \text{ and } 1.50$.

The BDS statistic is an "omnibus" statistic in the sense that it has good power properties against a variety of alternatives. It is not designed with any specific parametric alternative in mind. See Brock, Hsieh, and LeBaron (1991). The correlation dimension has its root in chaos theory (see Grassberger and Procaccia(1983)).

When we apply the BDS statistic to the residuals $\hat{I}(t)$ of our symmetric exchange rate decompositions we find values much larger than the α -critical value for the significance level

$\alpha = 0.001$. For $\hat{I}(t)$ based on 91 and 7 day trends, the values of the standardized BDS statistic (6) with $m=2, 3, 4, 5$, and $k=1$ are listed in Table 17. From Table 17, a hypothesis of i.i.d. is rejected. (These values are all between -1.96 and 1.96. Note that the upper 2.5% point of the standard normal distribution is 1.96.)

	JY	GM	SF	BP
m=2	11.230	7.925	9.139	10.900
	6.501	7.398	6.817	8.071
m=3	6.456	5.035	5.437	6.577
	4.018	4.956	4.561	4.867
m=4	5.742	4.598	4.837	5.952
	3.603	4.692	4.229	4.267
m=5	5.938	4.859	5.073	6.259
	3.703	4.929	4.373	4.383

< Table 17 >

Table 17. Values of the BDS statistic for $\hat{I}(t)$ of the (91,7) symmetric decomposition. In this table, $k=1$.

Note that all of the BDS statistics in this paper are calculated using 'data segment #1', which we picked from day 197 to day 696 of the data series, and 'data segment #2', from day 697 to day 1196. Thus the number in each segment is equally $T=500$.

The results of each segment are shown in the tables as follows: for example, the BDS statistic 11.230 in the first row of 'm=2' in the above table was calculated using data segment #1, on the other hand, for the case of data segment #2, when $m=2$, the BDS statistic is 6.501 as shown in the second row of 'm=2'. Or we will show the results such as 'BDS statistic of data segment #1' : 'BDS statistic of data segment #2' when the statistic are not listed on the table.

III-2. Autoregressive Decomposition of the Residuals $\hat{I}(t)$: Symmetric Decomposition.

We investigated, in Section II-1(v), the dependence of $\hat{I}(t)$ on its lagged values by using a multivariate autoregressive fit program ("ar" in S-PLUS). We found that the Akaike Information Criteria selects an order 4 autoregressive model. To check the success of the autoregressive fit we applied the BDS test to the residuals and found the values below .

	JY	GM	SF	BP
m=2	1.021	0.736	0.502	0.469
	1.032	0.590	0.732	0.576
m=3	0.923	0.393	0.259	0.254
	0.730	0.400	0.369	0.328
m=4	0.987	0.329	0.313	0.234
	0.760	0.351	0.246	0.339
m=5	1.058	0.310	0.307	0.320
	0.771	0.369	0.197	0.389

< Table 18-1 >

Table 18-1. Values of the BDS statistic for residuals of multivariate autoregressive fit in Section II-1(v). In this table, $k=1$.

Since the BDS statistic has considerable power when $T=500$, the values in the table indicate that the distribution of the residuals can be reasonably well approximated by the distribution of independent identically distributed random variables.

The values of BDS statistics for $\varepsilon_1(t)$, the residuals of univariate autoregressive fit of $I(t)$ for Japanese Yen and German Marks in Section II-1(v) are shown in Table 18-2. The values of the BDS statistics suggest that null hypothesis of i.i.d. can not be rejected.

	JY	GM
m=2	0.856	0.755
	0.883	0.564
m=3	0.850	0.397
	0.625	0.404
m=4	0.877	0.341
	0.657	0.380
m=5	0.935	0.297
	0.665	0.385

< Table 18-2 >

Table 18-2. Values of the BDS statistic for residuals of univariate autoregressive fit of $I(t)$ in Section II-1(v). In this table, $k=1$.

III-3. Predictive Decomposition

We also did an analysis of the predictive smoothers. The values of the BDS statistic when applied to the residuals $\hat{I}_p(t)$ of (91,7) predictive decomposition are shown in Table 19. The values of BDS statistic of $\hat{I}_p(t)$ equal -0.221:-0.093 and -0.238:-0.432 for Japanese Yen and German Marks when $m=2$, while these values are 0.282:0.056 and 0.039:-0.097 when $m=3$. This suggests the very parsimonious model $Y(t) = L_p(t) + S_p(t) + \varepsilon(t)$ with $\varepsilon(t)$ i.i.d. for Japanese Yen and German Marks. Since the value of BDS statistics are close to zero, the null hypothesis of i.i.d. can not be rejected. So we did not go further to see the autoregressive structure of $\hat{I}_p(t)$.

	JY	GM	SF	BP
m=2	-0.221	-0.238	-0.117	-0.104
	-0.093	-0.432	-0.404	-0.331
m=3	0.282	0.039	0.080	-0.024
	0.056	-0.097	-0.139	0.042
m=4	0.395	0.234	0.214	0.034
	0.208	0.039	-0.003	0.176
m=5	0.417	0.300	0.219	0.060
	0.288	0.042	-0.051	0.200

< Table 19 >

Table 19. Values of the BDS statistic for $\hat{I}_p(t)$ of the (91,7) predictive decomposition. In this table, $k=1$.

Chapter IV. Volatility Measures and Their Time Series Decompositions

The predictive decomposition of Section II-2 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 Section II, III, and IV.

A. Volatility Measures Based on Returns

Many commonly used measures of volatility are based on the conditional variance $\sigma^2(t)$ given the past of the returns

$$X(t) = \log \left[\frac{Y(t)}{Y(t-1)} \right]$$

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 $\sigma^2(t) = \text{Var}(X(t)|\Omega_{t-1}) = E[V(t)|\Omega_{t-1}]$. Thus we can think of $\sigma^2(t)$ as the conditional mean of $V(t)$, and apply the predictive decomposition of Section II-2. To illustrate, we take Ω_{t-1} to be the set of lagged values $X(t-1), X(t-2), \dots$ 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

$$m(t) = \frac{1}{t-1} \sum_{j=1}^{t-1} X(j)$$

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 II-2-(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 towards 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 Japanese Yen, there were only one missing data day, while for German Marks and Swiss Francs there were 2 and 2 missing data days. For these days, we used the average of the returns of the two closest days were 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 were 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 13 shows a graphical analysis of the volatility of Japanese Yen. The A part compares the 91 and 181 day long term trends $\hat{\sigma}_L(t)$ and changes in $V(t)$ while the 181 day trend make more moderate adjustments as $V(t)$ changes. The B part 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 14 gives a comparison of the volatilities of German Marks and Swiss Francs. They resemble each other, but the Swiss Francs are generally more volatile.

Figure 13-A : Standard Deviation Estimator : Long Term Trends (91 days and 181 days) : $\text{Log}Y(t) - \text{Log}(Y(t-1)) : \text{JY}$

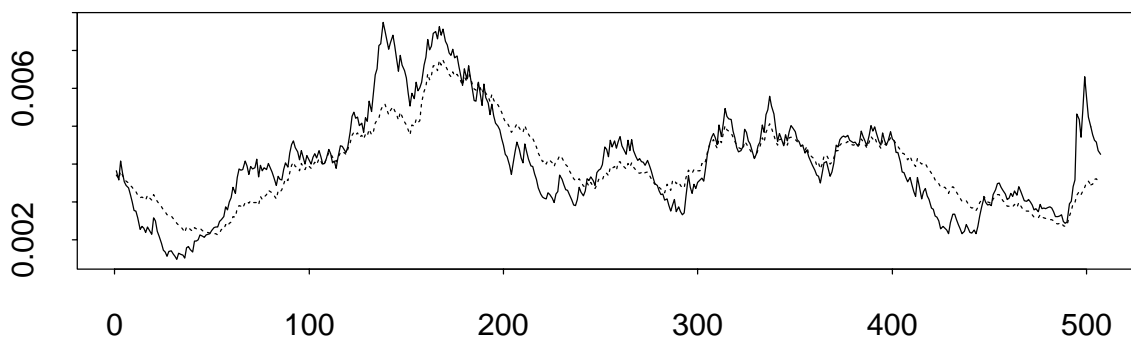


Figure 13-B : Standard Deviation Estimator : Long Term Trend and Long + Short Term Trend (181 days...15 day) : $\text{Log}Y(t) - \text{Log}(Y(t-1)) : \text{JY}$

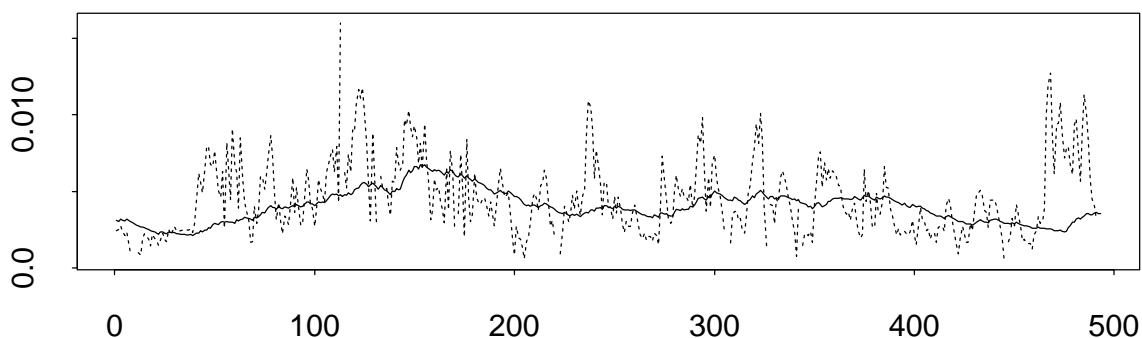


Figure 13. 13-A gives the long term volatility of Japanese Yen using predictive decomposition based on 91 ———, and 181 - - - - - day time intervals. 13-B gives the volatility of Japanese Yen using (181,15) predictive decomposition. The solid line is the long term trend while the dotted line is the sum of the long and short term trends.

Figure 14-A : Standard Deviation Estimator : Long Term Trends (181 days) :
 $\text{Log}Y(t) - \text{Log}(Y(t-1))$: GM and SF

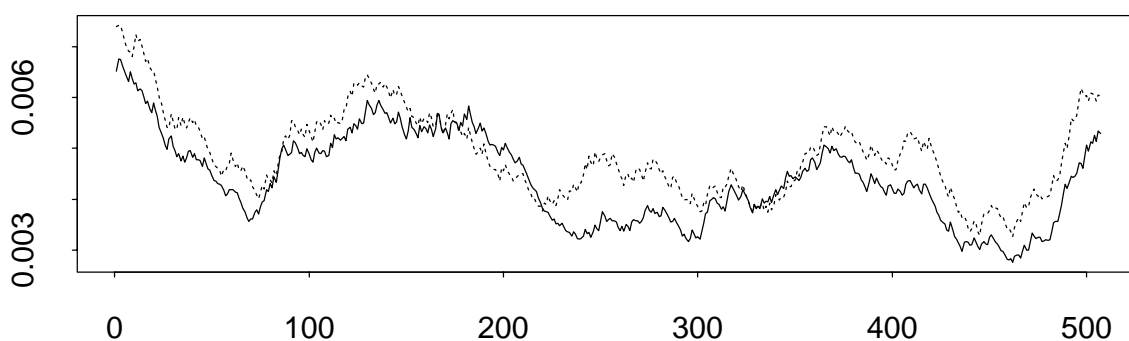


Figure 14-B : Variance Estimator : Short Term Trends (181 days...15 days) :
 $\text{Log}Y(t) - \text{Log}(Y(t-1))$: GM and SF

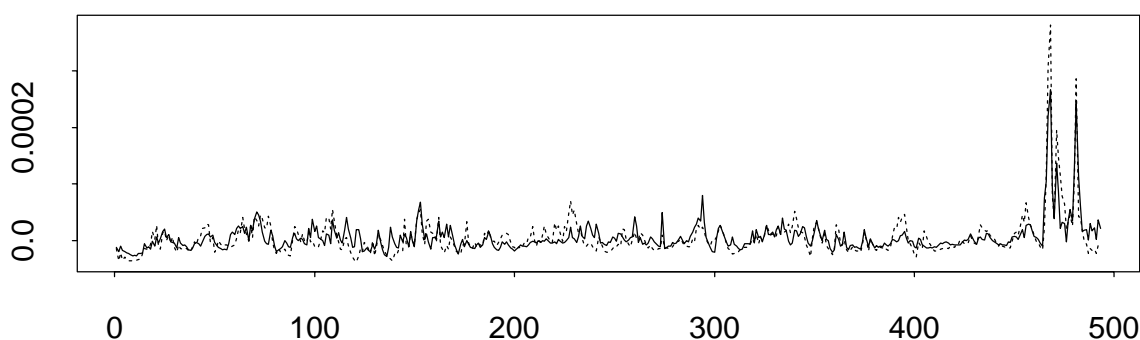


Figure 14. 14-A gives the long term volatility of German Marks ——— and Swiss Francs ······. The estimators are 181 day predictive smoothers. 14-B gives the short term trends of German Marks ——— and Swiss Francs ······ based on the (181,15) predictive decomposition.

B. RiskMetrics™

The J.P.Morgan RiskMetrics™ (1995) measure of volatility is an example of a predictive long term trend kernel estimator. In the third edition of RiskMetrics™ - Technical Document (page 80) they give

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

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

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

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

Figure 15-A : Percent of Changes : $\text{Log}Y(t) - \text{Log}(Y(t-1))$: JY

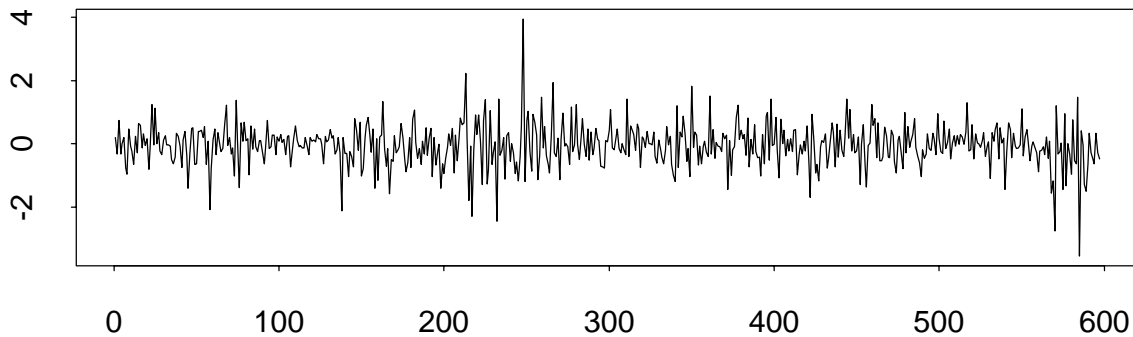


Figure 15-B : Standard Deviation Estimator : Predictive (91 days) and JPM (Approximation) : $\text{Log}Y(t) - \text{Log}(Y(t-1))$: JY

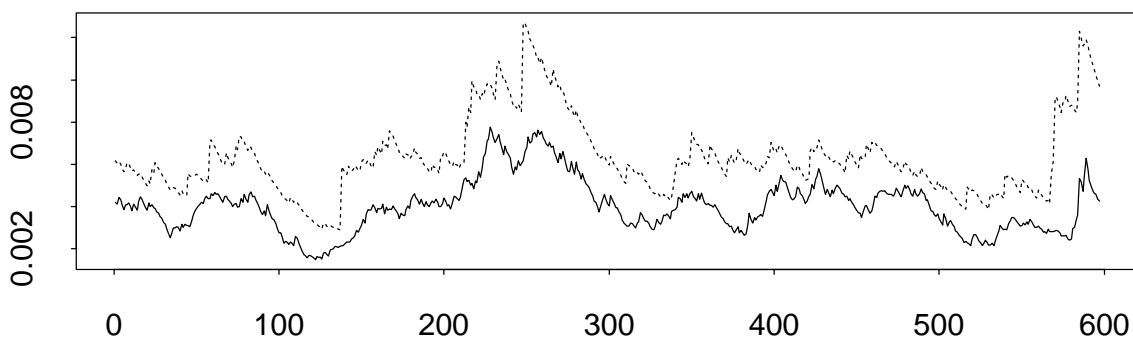


Figure 15. 15-A gives the returns of the Japanese exchange rate series. 15-B gives the 91 day long term trend volatility using our predictive decomposition — and the J.P.Morgan's RiskMetrics™ volatility of Japanese Yen using Tuesday-Friday data from the time period from January 6, 1992 to April 14, 1995.

Figure 15 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™ exponential kernel estimator. Both are based on using the first 91 days as historical data. The figure shows that both estimators give similar assessments of volatility with the 91 days predictive smoother having slightly less abrupt changes. In the appendix we argue that the locally linear predictive smoother shown in Figure 15, falls closer to the true volatility than RiskMetrics™. Note that RiskMetrics™ is higher than $\hat{L}_p(t)$ because it is based on using $\hat{w}(t) \equiv 0$.

C. Volatility Based on the Logarithmic Range

Next we turn to measures of volatility based on the logarithmic range, that is on the logarithm

$$W(t) = \text{Log} \frac{\text{High}}{\text{Low}} .$$

For this measure we will have 5 weekday values available, in our case a total of 860 days. Missing values were replaced by the average of the nearest available W values, while the W values for days with volume less than 200 were replaced by the average of that day's W score and the two nearest available W values.

Advantages of volatility measures based on the logarithmic range over measures based on returns have been discussed by Parkinson (1980), Garman and Klass(1980), Kupiec (1990), Kunitomo (1992), and Hsieh (1993), among others. Hsieh (1993) shows how logarithmic range volatility measures can be readily applied to risk analysis.

Figure 16-A : Standard Deviation Estimator : Long Term Trends (91 days and 181 days) :

$$\text{Log} \frac{\text{High}}{\text{Low}} : \text{JY}$$

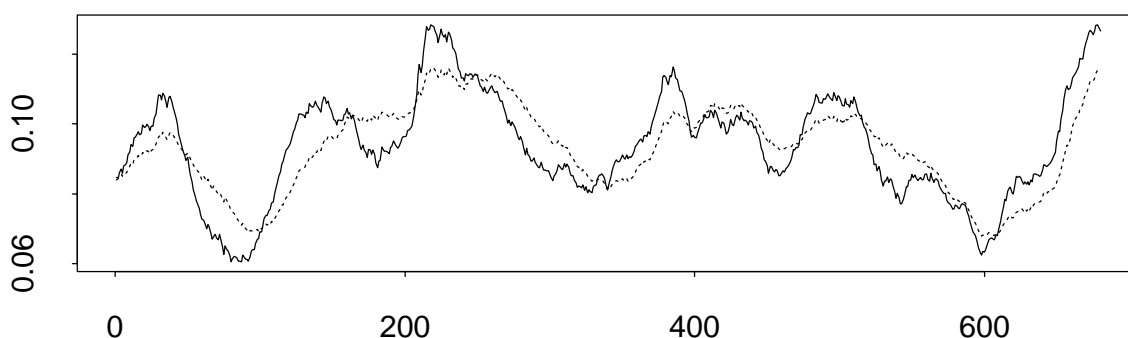


Figure 16-B : Standard Deviation Estimator : Long Term Trend and Long + Short Term Trend

$$(181 \text{ days } \dots 15 \text{ day}) : \text{Log} \frac{\text{High}}{\text{Low}} : \text{JY}$$

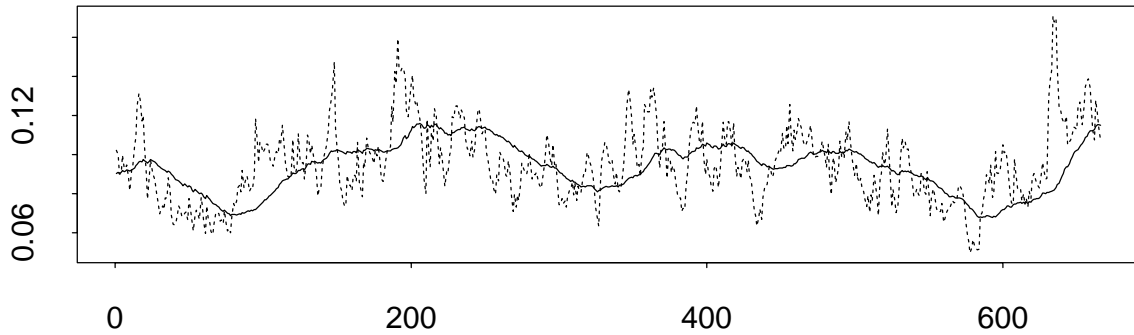


Figure 16. 16-A gives the long term logarithmic range volatility of Japanese Yen using LOWESS based on 91 ——— and 181 ······ day time intervals. 16-B gives the long term ——— and the long term plus short term ······ volatility based on the (181,15) decomposition.

Figure 17-A : Standard Deviation Estimator : Long Term Trends (181 days) : $\text{Log} \frac{\text{High}}{\text{Low}}$:

GM and SF

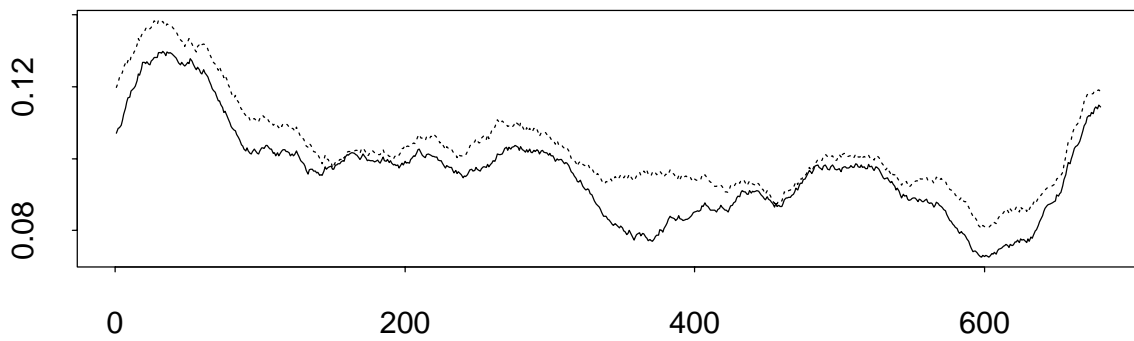


Figure 17-B : Short Term Trends (181 days ... 15 days) of the within day volatility of GM and

$$\text{SF} : \text{Log} \frac{\text{High}}{\text{Low}} : \text{GM and SF}$$

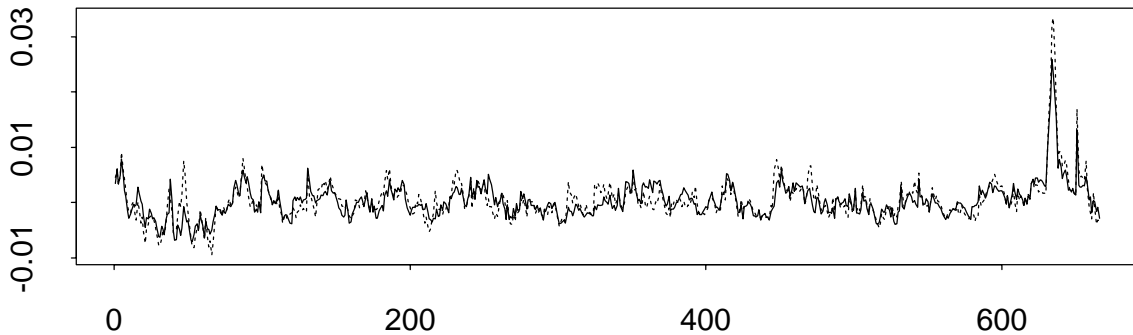


Figure 17. 17-A gives the long term logarithmic range volatility of German Marks — and Swiss Francs based on the LOWESS 181 day smoother. 17-B gives the short term trends of German Marks — and Swiss Francs based on the (181,15) decompositions.

Figure 16 shows the 91 day and 181 day long term volatility- trends $L_{PW}(t)$ based on $W(t)$. By comparing Figures 13 and 16 we see interesting similarities and differences between the volatilities based on returns and those based on the logarithmic range. This comparison sheds light on the differences between “between day volatility” and “within day volatility”. Figure 17 compares the volatilities of German Marks and Swiss Francs and shows similar tendencies to those in Figure 14.

Appendix. Asymptotic Approximations

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 leads 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™, 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™. We rescale by setting

$$u = \frac{k}{T+1}, v = \frac{t}{T+1}, h = \frac{M}{T}. \quad (8)$$

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

We define $\mu_{T^{-1}}(u) = \mu(\lfloor u(T+1) \rfloor)$, $u \in (0, 1)$, where $\lfloor \cdot \rfloor$ denotes the greatest integer function and assume that $\mu_{T^{-1}}(u) \rightarrow \mu_0(u)$, $u \in (0, 1)$, for some function $\mu_0(u)$ on $(0, 1)$. In other words, $\mu_0(u)$ is the mean function after the transformation (8) of the time scale to $(0, 1)$.

A. The Symmetric Regression Smoothers

In the case of the symmetric long term smoothing kernel weights of Section II-1-(i), we write $[-M_L, M_L]$ for the interval of time points k that the kernel is nonzero, and we set $h_L = \frac{M_L}{T}$. To use a very long term smoother is equivalent to assuming that for some $0 < h_0 < 1$, $h_L \rightarrow h_0$ as $T \rightarrow \infty$. We let $L_{T^{-1}}(u) = L(\lfloor u(T+1) \rfloor)$ be the long term trend in the new units where L is based on the weights (1), then

$$\lim_{T \rightarrow \infty} L_{T^{-1}}(v) = h_0^{-1} \int_{v-h_0}^{v+h_0} \mu_0(u) K\left(\frac{u-v}{h_0}\right) du = \int_{-1}^1 \mu_0(v+h_0x) K(x) dx. \quad (9)$$

where $K(\cdot)$ is a kernel function used in the Section II-1.

We call that the limit $L_0(v)$ in (9) the *asymptotic long term trend*.

For the short term kernel weights of Section II-1-(ii) we introduce $h_s = \frac{M_s}{T}$ 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 $h_s \rightarrow 0$ as $T \rightarrow \infty$. Let $S_{T^{-1}}(u) = S([u(T+1)])$ be the short term trend in the $(0, 1)$ units. Since $h_s \rightarrow 0$, we can obtain useful approximations by expanding $S_{T^{-1}}(v)$ in a Taylor series :

$$\begin{aligned} S_{T^{-1}}(v) &= C^{-1} \sum_u K\left(\frac{u-v}{h_s}\right) [\mu_0(u) - L_0(u)] \\ &= (h_s T)^{-1} \sum_u K\left(\frac{u-v}{h_s}\right) \left[\mu_0(v) - L_0(v) + [\mu_0'(v) - L_0'(v)](u-v) \right. \\ &\quad \left. + \frac{1}{2} [\mu_0''(v) - L_0''(v)](u-v)^2 \right] + O(h_s^3) \end{aligned} \quad (10)$$

where we have assumed that $\mu_0(u)$ has a continuous third derivative at v and that $Th_s^3 = \frac{M_s^3}{T^2} \rightarrow \infty$ as $T \rightarrow \infty$.

Since the kernel K is symmetric, the derivative term in the expansion (10) is zero. By setting $x = \frac{u-v}{h_s}$, we find the following approximation $S_0(v)$ to the short term trend

$$S_0(v) = \mu_0(v) - L_0(v) + \frac{1}{2} [\mu_0''(v) - L_0''(v)] h_s^2 \sigma_K^2 + O(h_s^3) \quad (11)$$

where $\sigma_K^2 = \int_{-1}^1 x^2 K(x) dx$ is the variance of K . This approximation shows how the short term trend depends on the instantaneous value $\mu_0(v) - L_0(v)$, the second derivative, the band width 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 $h_L \rightarrow 0$ and $(h_s/h_L) \rightarrow 0$. In this case the trend considered is moderately long. From (9) we find that the moderately long term trend is

$$L_0(v) = \mu_0(v) + \frac{1}{2} \mu_0''(v) h_L^2 \sigma_K^2 + O(h_L^3) + O(h_s^2).$$

By combining this with (11) we find

$$S_0(v) = -\frac{1}{2} \mu_0''(v) h_L^2 \sigma_K^2 + O(h_L^3).$$

This approximation shows that the second derivative has a crucial influence on the short 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).

B. The Predictive Regression Smoothers

If we apply the preceding limits and approximations to the predictive smoothers in Section II-2 based on weights of the form (1) we find the long term approximations

$$L_{op}(v) = 2h_0^{-1} \int_{v-h_0}^v \mu_0(u) K\left(\frac{u-v}{h_0}\right) du = 2 \int_{-1}^0 \mu_0(v+h_0x) K(x) dx, \quad 0 < h_0 < v$$

$$S_{op}(v) = \mu_0(v) - L_{op}(v) + [\mu_0'(v) - L_{op}'(v)] h_S \mu_{PK} + \frac{1}{2} [\mu_0''(v) - L_{op}''(v)] h_S^2 \mu_{PK}^2 + O(h_S^3)$$

where $\mu_{PK} = 2 \int_{-1}^0 x K(x) dx$ and $\sigma_{PK}^2 = 4 \int_{-1}^0 x^2 K(x) dx$.

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-1 do not have this property. It is possible to show that for the locally linear short term predictive smoother based on weights of the form (4)

$$S_{op}(v) = \mu_0(v) - L_{op}(v) + \frac{1}{2} [\mu_0''(v) - L_{op}''(v)] h_S^2 \mu_{PK}^2 + O(h_S^3).$$

From this it follows that 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 $h_L \rightarrow 0$ and $(h_S/h_L) \rightarrow 0$, we obtain, for smoothers based on the weights (1), the approximations

$$L_{op}(v) = \mu_0(v) + \mu_0'(v) h_L \mu_{PK} + \frac{1}{2} \mu_0''(v) h_L^2 \mu_{PK}^2 + O(h_L^3)$$

$$S_{op}(v) = \mu_0'(v) h_L \mu_{PK} + O(h_L^2)$$

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, Massachusetts : 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, 1997.
- 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, "Calendar Effects in Monetary Time Series: Modeling and Adjustment," *Journal of American Statistical Association*, 77, 1982, pp. 520-528.
- 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 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 : John Wiley & Sons, Inc., 1994.
- Engle, Robert E., "Autoregressive Conditional Heteroskedasticity," *Econometrica*, Vol. 50, 1982, pp. 987-1007.
- Engle, Robert F. and W.J. Granger, "Co-Integration and Error Correction : Representation, Estimation and Testing," *Econometrica*, Vol. 55, No. 2, March 1987, pp. 251-276.
- Fan, Jianqing, and Irène Gijbels, "Local Polynomial Modeling and its Applications," London: Chapman and Hall, 1995.
- Fan, J. "Local linear regression smoothers and their minimax efficiency." *Annals of Statistics*. Vol.21.1993. pp.196-216.
- Garman, Mark B. and Michael J. Klass, "On the Estimation of security price volatilities from Historical Data," *Journal of Business*, Vol. 53, No. 1, 1980, pp. 67-78.

- 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.
- Kunitomo, Naoto, "Improving the Parkinson Method of Estimating Security Price Volatilities," *Journal of Business*, Vol. 65, No. 2, April 1992, pp. 295-302.
- Kupiec, P., "Futures, Margins and Stock Price Volatility : Is There Any Link? Board of Governors of the Federal Reserve System," Finance and Economics Discussion Paper, No. 104, 1990.
- Masry, Elias and Dag Tjøstheim, "Nonparametric Estimation and Identification of Nonlinear Arch Time series : Strong Convergence and Asymptotic Normality," *Econometric Theory*, Vol. 11, 1995, pp. 258-289.
- Miura, Ryoza and H. Kishino, "Pricing of Bonds and their Derivatives with Multi-factor Stochastic Interest Rates : A Note," *Nonlinear and Convex Analysis in Economic Theory. Lecture Notes in Economics and Mathematical Systems*, pp. 419, Springer-Verlag, 1995.
- Morgan Guaranty Trust Company, "RiskMetrics™ . Technical Document," New York: 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.
- Parkinson, M., "The Extreme Value Method of Estimating the Variance of the Rate of Return," *Journal of Business*, 53, 1980, pp. 61-65.
- Phillips, P. C. B., "Optimal Inference in Cointegrated Systems," *Econometrica*, Vol. 59, No. 2, March 1991, pp. 283-306.
- Ruppert, D. and Wand, M.P., "Multivariate Locally Weighted Least Squares Regression," *The Annals of Statistics*, Vol. 22, No. 3, 1994, pp. 1346-1370.
- Shibata, Ritei, and Miura, Ryoza, "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.