Decomposition of Japanese Yen Interest Rate Data through Local Regression

By

Ritei Shibata and Ryozo Miura

Technical Report No.434
July 1995

Department of Statistics
University of California
Berkeley, California

DECOMPOSITION OF JAPANESE YEN INTEREST RATE DATA THROUGH LOCAL REGRESSION

RITEI SHIBATA AND RYOZO MIURA

Abstract. Seven different Japanese Yen interest rates recorded on a daily basis for the period from 1986 to 1992 are simultaneously analyzed. By introducing a new concept of "short term trend", we decompose each interest rate series into three components, "long term trend", "short term trend" and "irregular" by a two step *lowess* smoothing procedure. Furthermore, a multivariate autoregressive model (MAR) is fitted to the seven irregular series. The decomposition and the model fitting were quite satisfactory, and each component and the residuals of the MAR model are statistically well behaved. Thus it enables us to understand well various aspects of interest rate series from those trends, the MAR(2) coefficients, and its residuals. The result is compared with the decomposition through *sabl* and the advantages of our procedure will be discussed in relations to other parametric model fitting like ARCH or GARCH. Based on the decomposition we can have better daily prediction and more stable long term forecasting.

1. Introduction

In this paper, we are concerned with simultaneous analysis of seven different Japanese Yen interest rates recorded daily for the period beginning from the 1st of December, 1986 to the 16th of September, 1992. The basic idea behind our analysis is to decompose each time series into three components, namely "long term trend", "short term trend" and "irregular". The long term trend is obtained by applying a smoothing technique, a local regression procedure *lowess* or *loess* (Cleveland and Devlin, 1988) to the original time series with a yearly span, and the short term trend is obtained by applying again the *lowess* with a monthly span to the residuals of the smoothing. The decomposition explains well the behavior of the underlying structure of these seven interest rate series. For example, the behavior of the long term trend significantly links to that of the Japanese official discount rate and explains well the difference between two groups; a group of Euro interest rate and the other group of LIBOR(London Interbank Offered Rate) swap interest rate series, particularly well

Key words and phrases. Decomposition of Time Series, Local Regression, Short Term Trend, Sabl, Smoothing, Yen Interest Rates.

This work was partly done during the time when the first author was staying at Department of Statistics, U.C. Berkeley and the second author was at Department of Statistics, Harvard University.

when Japan is in a period of "bubble economy". The short term trend indicates a cycle though not exact. More importantly, a multivariate autoregressive model of order 2 fits well to the vector of seven irregular series and a good prediction of each interest rate is available on daily basis. It is enough to predict the irregular series based on the model, since daily changes of the long and short term trends are negligible relative to those of the irregular series.

Our approach is not based either on any definite economic theory or on financial theory, but based on a modern technique of data analysis, a nonparametric smoothing. In the latter part of this paper the validity of the decomposition will be checked from various statistical points of view: goodness of fit, accuracy of prediction, and the stability of the results, as well as the interpretability of the results. There still remain a lot of statistical properties to be checked, for example, a check of the independence of residuals besides the orthogonality. We leave such interesting problems for further investigation.

2. Data

Data we are analyzing are 3 month, 6 month and 1 year Euro Yen interest rate series, and 3 year, 5 year, 7 year, and 10 year LIBOR swap rate series. All series are daily for the period; from the 1st of December, 1986 to the 16th of September, 1992. This data is now regarded as historic in the sense that the period of over heated Japanese economy, or the "bubble" economy, is included. The length of each series is 2115, so that this study requires a powerful computing environment since computer intensive smoothing techniques like *sabl* or *lowess* is used.

Euro Yen interest rate is a fixed annual interest rate for Yen bonds, which are mainly traded in European countries for short term maturities, such as 3 months, 6 months or 1 year. On the other hand, LIBOR swap rate is a fixed annual interest rate with which the coupon can be swapped for a LIBOR 6 month variable interest rate contract for the same period. Swapped is only a coupon at a time, not a principal. Since there is a difference between the offer and the bid swap rates, which is mostly ranging from 0.002 % to 0.004 %, we analyze the mean of those two rates, so that meaningful scale unit of rate changes in practice would be around ± 0.004 %. It is also said that LIBOR swap rate reflects a long term forecast of economy and Euro Yen rate reflects various short term expectations.

Preceding to the analysis, we interpolated the rates for Sunday or other holidays by a simple linear interpolation method. Also the data for the 29th of February of a leap year are all removed to adjust the number of days in a year to 365 days. The details will be found in the last section.

3. Decomposition

3.1. Sabl decomposition. As a preliminary analysis, we tried the *sabl* (Seasonal Adjustment at Bell Laboratories) decomposition procedure which is proposed by

Cleveland and Devlin(1988) and implemented as a function `sabl` in S (Becker et al., 1988). This *sabl* procedure is widely used as well as X-11 for time series analysis of real data. By this procedure, the given time series is decomposed into three parts,

$$\text{original series} = \text{trend} + \text{seasonal} + \text{irregular}.$$

The decomposition is the result of a repeated application of smoothings to obtain the trend and seasonal components in turn (Cleveland et al., 1981). The initial (temporary) trend $T_0(t)$ is obtained from the original $Z(t)$ by applying a moving median smoothing with the given span and then two times moving average smoothings. The first moving average is done with the given span and the second is done with three time point span. The first seasonal component $S_1(t)$ is then obtained by applying a weighted moving median regression and a moving least squares regression to the moving subseries of residuals $Z(t) - T_0(t)$ with the given span. The next trend $T_1(t)$ is obtained in the same way from the residuals $Z(t) - S_1(t)$. A robustness weight is computed from the residuals $Z(t) - T_1(t) - S_1(t)$, and the second seasonal series $S_2(t)$ is obtained by applying a weighted moving robust regression to the residuals $Z(t) - T_1(t)$. The second trend $T_2(t)$ is obtained from $Z(t) - S_2(t)$ in the same way by using the robustness weight computed from the residuals $Z(t) - T_1(t) - S_2(t)$. The final two steps are a repeat of the last two steps, and the final estimates $S_3(t)$ and $T_3(t)$ follow. The irregular series is the series of residuals, $Z(t) - T_3(t) - S_3(t)$.

An important feature of this procedure is to be able to decrease the end effects of the smoothing since the weighted moving robust regression is used from the second stage. The procedure also makes it possible to extrapolate the long and short term trends. Another feature of this procedure is to eliminate or reduce interaction between the seasonal and the trend series by applying a power transformation of the original series. The power p of the transformation is selected so as to minimize the absolute value of the regression coefficient θ of the model

$$g(Z(t)) = T(t) + S(t) + \theta(S(t) - \bar{S})(T(t) - \bar{T}) + I(t),\ t = 1, 2, ..., n,$$

where g is a power transformation

$$g(x) = \begin{cases} x^p & \text{if } p > 0\ , \\ \log(x) & \text{if } p = 0\ , \\ -x^p & \text{if } p < 0\ , \end{cases}$$

and the $\bar{S}$ and $\bar{T}$ is the mean of the obtained components $S(t)$ and $T(t)$ respectively.

The following Figure.1 and Figure.2 are parts of the results of *sabl* decomposition of the 3 month Euro Yen interest rates. Here we simply assumed that the seasonality is a year, 365 days so that the smoothing span to obtain a trend is taken to be the same 365 days for both cases. On the other hand the seasonal component is obtained by smoothing the residuals of the trend over the same days of the year for 3 and 7 neighbor years. They are shown in Figure.1 and Figure.2 respectively. More explicitly

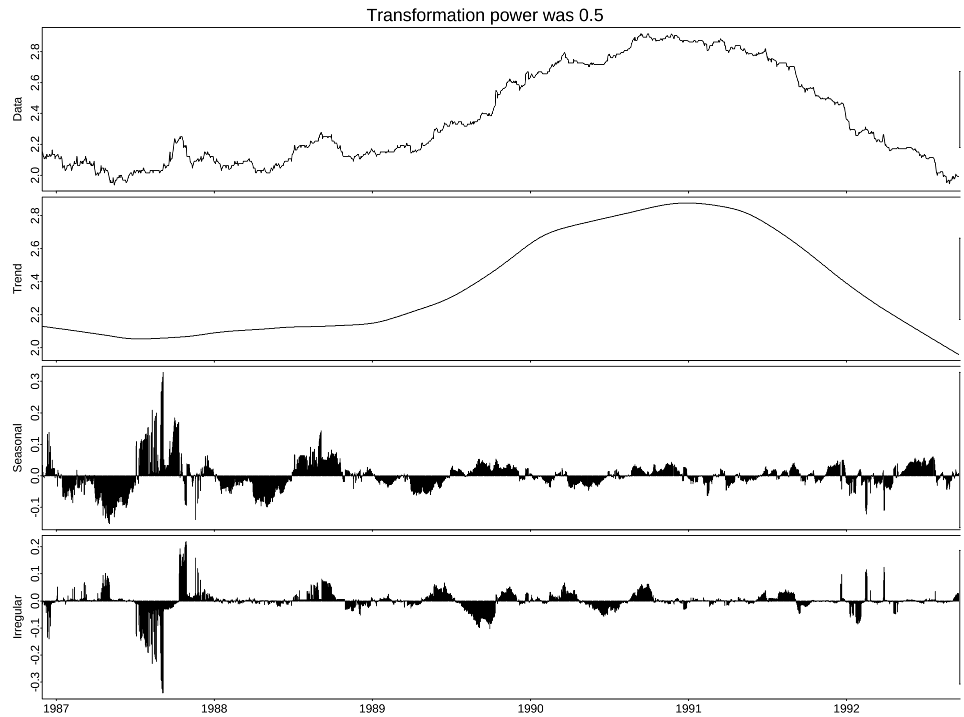

FIGURE 1. *sabl* decomposition of the 3 month Euro Yen interest rate series. The seasonal component on the third panel is obtained by smoothing over the same days in 3 neighbor years.

the argument `trend` of the S function `sabl` is taken to be 365 and an optional argument `seasonal` is taken to be 3 and 7 respectively. The data is organized as an S time series object with the attribute `frequency=365`. These figures are plotted by the S function `sablplot`, As is shown in those figures, the power of the transformation selected by the *sabl* procedure is $p = 0.5$, that is, this procedure concluded that the square root transformation is appropriate for eliminating interaction between the trend and seasonal. In each of the two figures the four panels show the square root transformed original series, the trend, the seasonal and the irregular, from the top to the bottom. The bar at the right side of each panel is to indicate the unit scale of each panel. It shows the same length in the different scales across the panels. This makes the comparison by eyes easier.

The problem of such a decomposition is now clear. The irregular part behaves so as to compensate the cyclic behavior of the seasonal series. It does not look irregular nor random but another seasonal component. Such a result suggests that no significant yearly cycle exists in the series. We tried many other cycles, weekly, monthly, quarterly for all seven series, but could not find any definite seasonality.

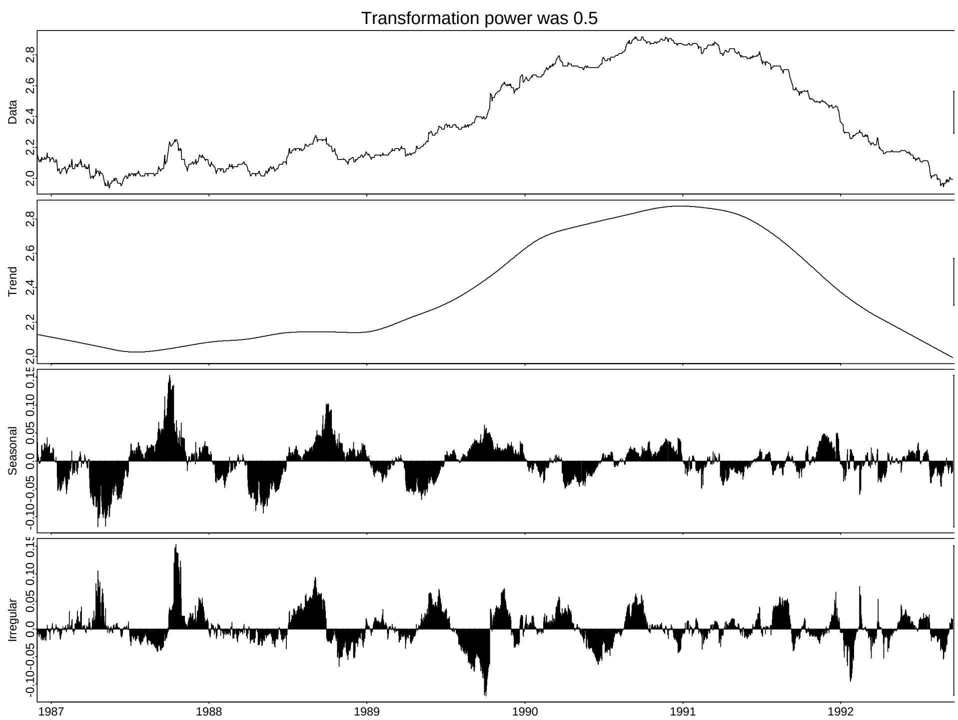

FIGURE 2. *sabl* decomposition of the 3 month Euro Yen interest rate series. The seasonal component on the third panel is obtained by smoothing over the same days in 7 neighbor years.

3.2. New decomposition. Based on the preliminary analysis, we came to a concept of "short term trend" in place of seasonality. The short term trend is not seasonal but moves very slowly on a monthly scale. In contrast, we then call a trend like that obtained by *sabl* as "long term trend", which changes slowly on a yearly scale. The concepts of "long term trend" and "short term trend" seem closely related to those of "permanent" and "transitory" components in econometrics, which are applied to the detection of "business cycle". For example, Beveridge and Nelson (1981) explicitly proposed such a decomposition by explicitly defining those two components (i.e. permanent and transitory) for an ARIMA(p, 1, q) process $\{z_t\}$. Their definition of permanent component is a process $\bar{z}_t$ which satisfies

$$\bar{z}_t - \bar{z}_{t-1} = \mu + \left(\sum_{j=0}^{\infty} a_j \right) \epsilon_t$$

and the transitory component as a process

$$c_t = - \left(\sum_{j=1}^{\infty} a_j \right) \epsilon_t - \left(\sum_{j=2}^{\infty} a_j \right) \epsilon_{t-1} - \cdots ,$$

observing that

$$\lim_{k \to \infty} \{ \mathrm{E} \left(z_{t+k} | z_t, z_{t-1}, \cdots \right) - k\mu \} = \bar{z}_t.$$

Here we assumed that the ARMA(p, 1, q) process $\{z_t\}$ is represented as

$$z_t - z_{t-1} = \mu + \sum_{j=0}^{\infty} a_j \epsilon_{t-j}$$

with innovations $\{\epsilon_t\}$ of the differences of $\{z_t\}$. Their idea is that the permanent component should be an essential part of a long-term forecast, that is, the conditional expectation of z_{t+k} at the time of t for very large k. They applied such a decomposition procedure to the various economic indices including GNP for the detection of business cycles of U.S. economy during the period of 1947 to 1977. A related work is Nelson and Plosser(1982). More recently Cochrane (1988) developed an estimation procedure for the variance of the permanent component. A mathematical approach is done by Quah(1992). Also a good review of this approach can be found in Enders(1995).

However, our approach is in fact different from their approach. First of all, we don't assume any particular model like ARIMA behind our analysis, and the long term and the short term trends are both assumed to be non stochastic, rather deterministic. And, we decompose the original time series into three components, "long term trend", "short term trend" and "irregular". Conceptually we might say that the former two components are corresponding to their "permanent" component, and the last "irregular" to their "transitory" component. As is seen later, distinction between long term and short term changes of interest rates are quite important in analyzing such interest rate series not only for prediction also for detection of business cycles. Also we believe that our procedure is, in general, a more flexible vehicle to extract interesting features of the phenomena from a given data, than an automatic application of a particular model like ARCH or GARCH (See the special volume of Journal of Econometrics 52, 1992 for the modeling economic data).

To extract both the long and the short term trends, we use a smoothing technique called *lowess* proposed by Cleveland and Devlin(1982). With *lowess* the smoothed values are obtained by applying a local linear regression with the weights reciprocal to the distance in time of data points from the time point at which we are now smoothing. The weights are then made decreased for the data with large residuals, and again a local linear regression with the new weights is applied. This process continues as long as the smoothed value changes significantly. This smoothing technique is linear with respect to the data in each step but is not linear in the overall procedure since the weights are updated according to the value of the residuals. The algorithm we actually used is the function `lowess` which is implemented in the S software (Becker and Chambers , 1988).

The two smoothings are not commutative. We have applied two smoothings to obtain the long term trend at first and the short term trend at second. We have used a year span for smoothing the original data to get a long term trend and a month span for smoothing the residuals of the long term smoothing to get a short term trend. If we reverse the procedure we obtain something different. The short term trend includes the effects of long term trends. The reader may wonder if any other choice of smoothing span is available to get a short term trend. We tried several other choices including weekly, bimonthly, quarterly. Our criterion is whether we obtain the stationarity of irregular series and the interpretation of short term trend. Weekly span does not provide us a good interpretation of the short term trend and any other choices longer than a month give us almost the same short term trend but the irregular component looks like less stationary.

The following Figure.3 shows an example of our decomposition, the decomposition of the 3 month Euro Yen interest rate. Our notation for the decomposition of the ith interest rate series $Z_i(t)$ into the long term $L_i(t)$, short term $S_i(t)$, and irregular series $I_i(t)$ is

$$Z_i(t) = L_i(t) + S_i(t) + I_i(t),$$

for $i = 1, \cdots, 7$. An advantage of this new decomposition procedure to the previous *sabl* procedure is now clear. The irregular series we obtained does not exhibit any peculiar behavior and looks very stationary.

4. Three components of the decomposition

4.1. Long term trend. Seven long term trends are shown together in Figure.4. Each long term trend $L_i(t)$ is respectively obtained from the corresponding original

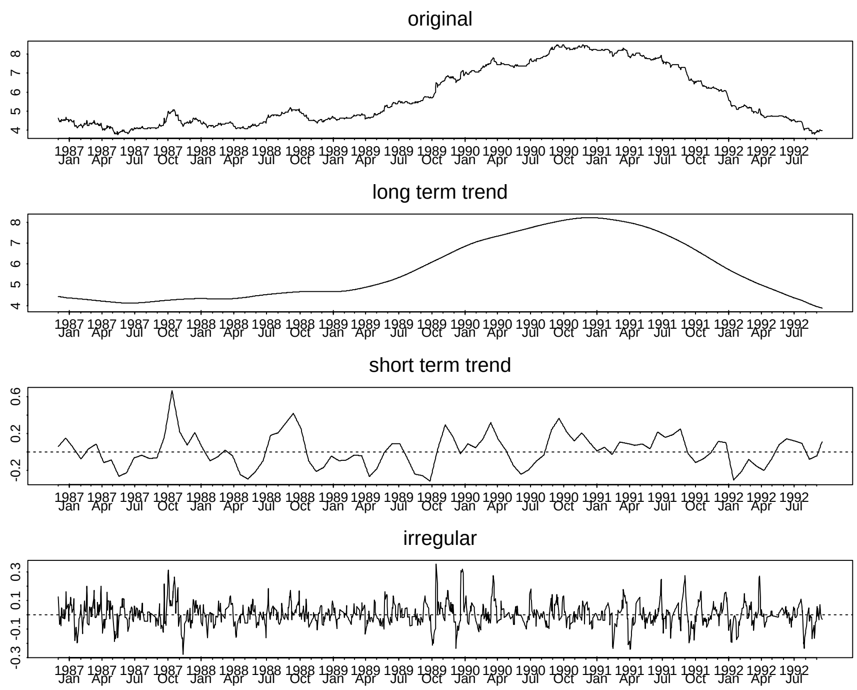

FIGURE 3. Decomposition of the 3 month Euro Yen interest rate series on the top panel into the following three components, the long term trend, the short term trend and the irregular series.

TABLE 1. Change Date of Japanese Official Discount Rate (%)

86-12-01	87-02-23	89-05-31	89-10-10	89-12-25	90-03-20
3	2.5	3.25	3.75	4.25	5.25

90-08-30	91-07-01	91-11-14	91-12-30	92-04-02	92-07-27
6	5.5	5	4.5	3.75	3.25

time series $Z_i(t)$ independently. The Japanese official discount rate given in Table.1 is also plotted in the figure for reference. In Table.1, the date of changes and the new rate are shown except that the first date 86-12-01 is not a change date but the date of the beginning of our series analyzed. The graph of each trend shown in Fig.4 naturally links with the official discount rate. The seven trend series are roughly classified into two groups, one for the 3 months to 1 year interest rates and the other for the 3 years to 10 years swap rates. As is well known, the two groups changed their relative position in the period of the "bubble" economy, around 1990 to 1992. The order of the magnitude of each trends in either group is quite natural, that is, upward sloping yield curve, until the beginning of the bubble economy. The 3 months, the 6 months and the 1 year in the first group are lower than the 3 years, the 5 years, the 7 years and the 10 years in the second group. During the bubble period, the order is totally reversed, that is, downward sloping yield curve, although the order is reversed except the 3 months in the former half of the period. An interesting observation is that the order of magnitude in the second group got back to the original one later in the year of 1991, but the order in the first group did not. Their order remains the same even after the bubble economy collapsed. This explains well the difference between the two groups. The rates in the first group are Euro Yen interest rate for short term contracts and those in the second group are LIBOR swap rate for long term contracts. Any economic interpretation is beyond this paper and left for future investigation.

We can check the smoothness of the long term trend with the boxplots of the lag 1 difference of the series shown in Figure.5. Again, the frequency distribution of the one day difference of long term trend can be classified into two groups, one for the 3 months to 1 year interest rates and the other for the 3 year to 10 year swap rates. It shows a greater volatility in the second group. We conclude the comments by mentioning that these seven long term trends show together not only the global movements but the differences of two groups both globally and locally.

4.2. Short term trend. The seven short term trends $S_i(t), t = 1, \cdots, 7$ are shown together in Figure.6. Apart from the long term trends, all the seven short term trends behave very similarly. The frequency distributions of the one day differences are also similar, as shown in Figure.7. The main effects caused by changes in the official discount rate and other economic environments are well removed by extracting the

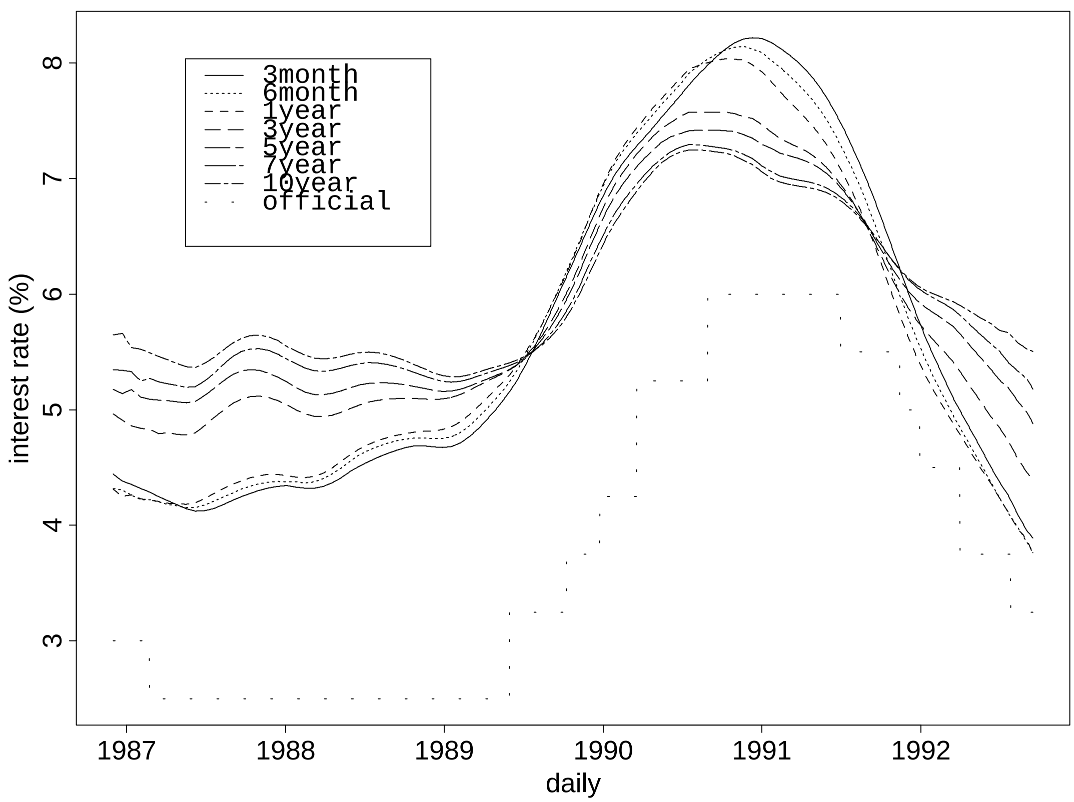

FIGURE 4. Long term trends and the official discount rate series.

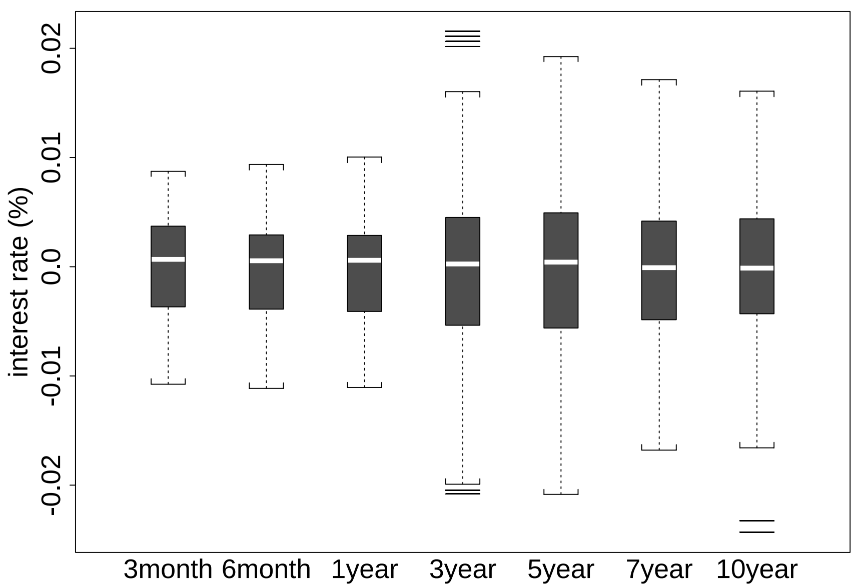

FIGURE 5. Boxplots of lag 1 differences of the long term trends.

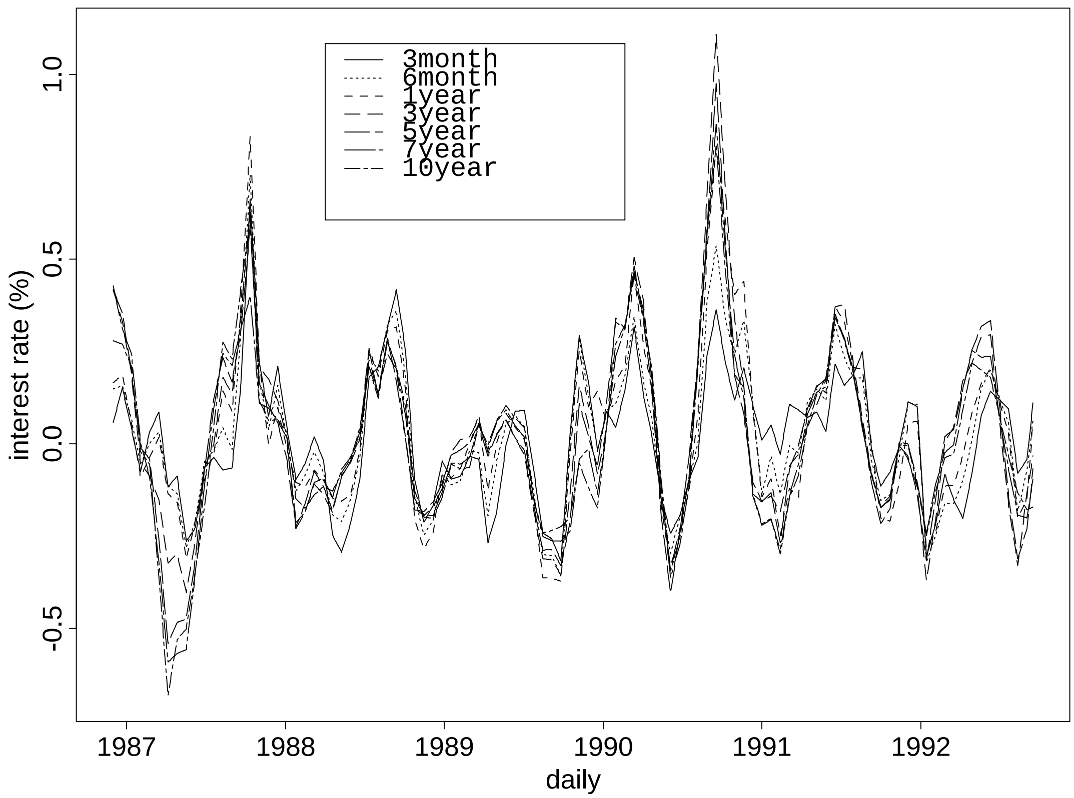

FIGURE 6. Short term trends

long term trend from the original series and hence no significant long term difference is left in the seven short term trends. We note that such a check is possible because we are not analyzing a single interest rate but seven different interest rate series simultaneously. An apparent fact we can see in Figure.6 is that all of the short term trends simultaneously increase from June and decreases around December every year. It is, of course, not a definite cycle nor a seasonality. To clarify this kind of business cycle, we discretized the trend. The discretization here means that each value is replaced by the mean of the values over the intervals where the series has the same sign all the time. This device makes such a behavior clearer as is seen in Figure.8. Table.2 is a list of the dates just after each short term trend crosses the level 0. The symbol $\searrow$ indicates a crossing from positive to negative, and the $\nearrow$ indicates that from negative to positive. The dates close to each other are grouped into a single row and the change dates of the official discount rate are also shown as a reference in the last column with the symbols which indicate whether it is increased or decreased. We see from this table that the dates of the crossings for those short term trends are quite close to each other and that there are fewer crossings for longer term interest

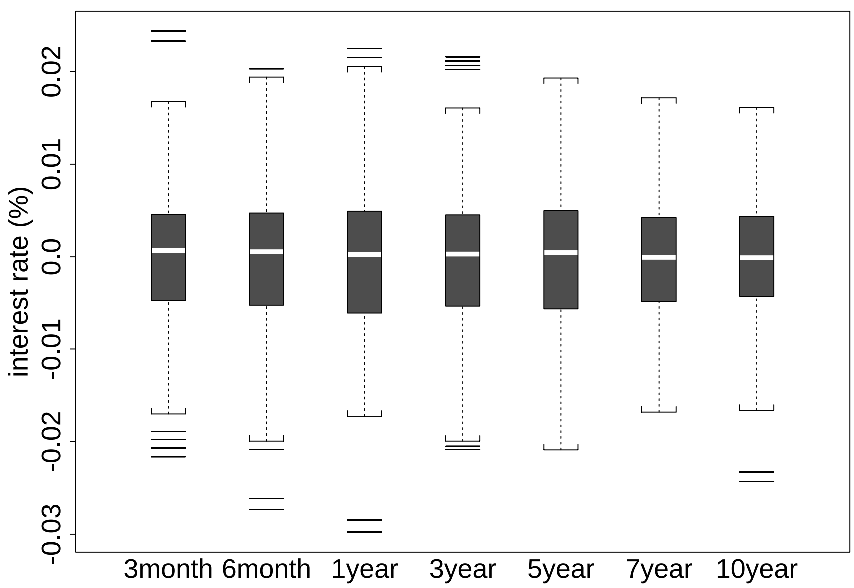

FIGURE 7. Boxplots of lag 1 differences of short term trends

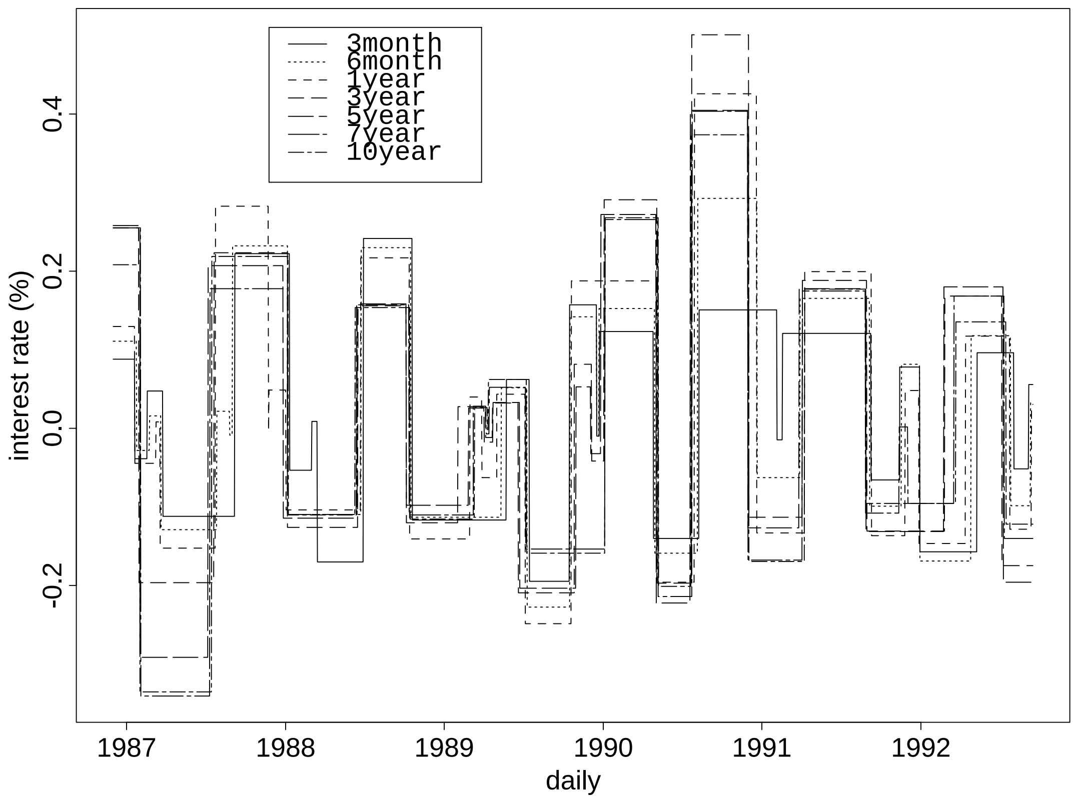

FIGURE 8. Discretized short term trends

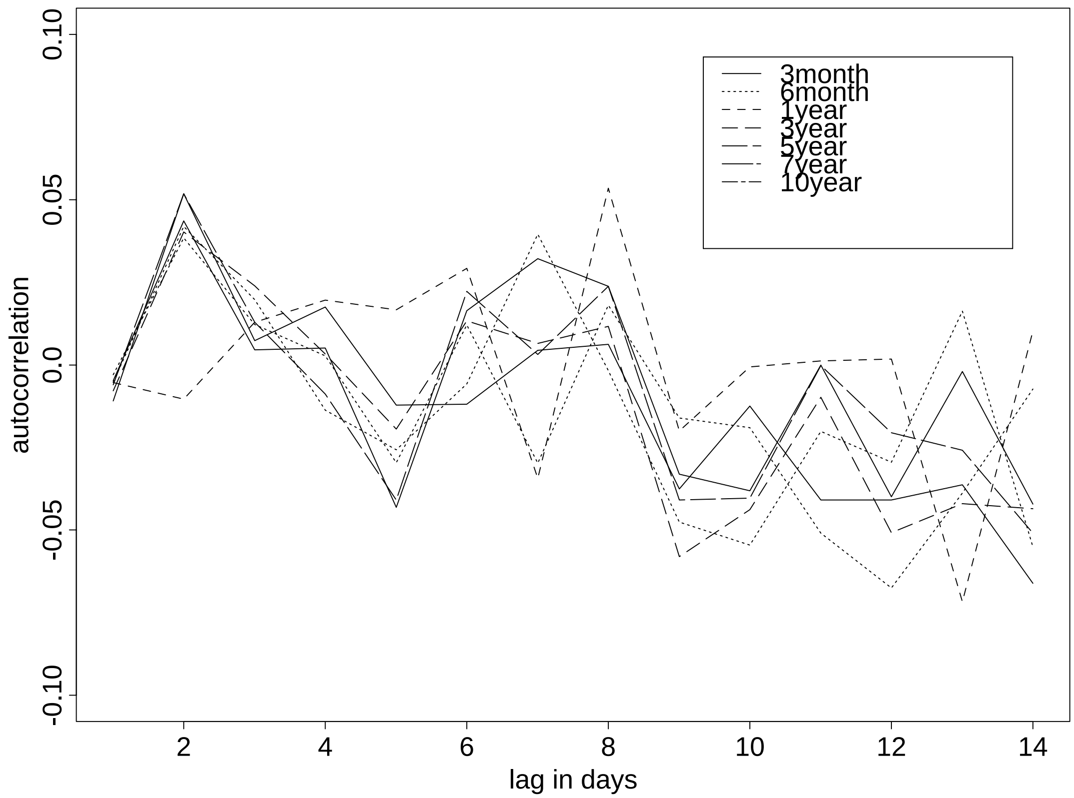

FIGURE 9. Autocorrelations of residuals

rates. Also we should note that the short term trends are not so related to the changes of the official discount rates. This indicates that the changes of the official discount rate mainly affects the long term trend, and thus justify the plausibility of our decomposition in this sense.

4.3. Irregular. The seven irregular series were combined into a multivariate time series $I(t) = (I_1(t), \cdots, I_7(t))^T$ and a multivariate autoregressive model MAR(2),

$$I(t) = A\,I(t-1) + B\,I(t-2) + \epsilon(t)$$

is fitted to. The order of the autoregression is selected so as to minimize AIC. Autocorrelations of the residuals $\epsilon_i(t), i = 1, \cdots, 7$ are shown in Figure.9. There is no doubt on the orthogonality of each residuals, since all autocorrelations are between -0.06 and 0.05. This supports that the fit of the MAR(2) model is quite good.

It is interesting to note that the coefficient matrix A in Table.3 shows that the 6 month rate significantly affects the 3 month rate and also affects other rates except the 10 years. This behavior is understandable since the volume of 6 month coupon

Table 2. Zero Crossing Date of Short Term Trends

	3 months	6 months	1 year	3 years	5 years	7 years	10 years	official
↘	87-01-20	87-01-24	87-01-20	87-01-29	87-02-01	87-02-03	87-02-01	
↗	87-02-17	87-02-23	87-03-10					↘ 87-02-23
↘	87-03-25	87-03-21	87-03-19					
↗		87-07-27	87-07-24	87-07-21	87-07-07	87-07-11	87-07-16	
↘		87-08-26						
↗	87-09-07	87-09-01						
↘			87-11-23					
↗			87-11-24					
↘	88-01-11	88-01-06	88-01-02	88-01-06	87-12-27	88-01-06	88-01-08	
↗	88-03-03							
↘	88-03-15							
↗	88-06-29	88-06-24	88-06-23	88-06-17	88-06-13	88-06-15	88-06-10	
↘	88-10-20	88-10-16	88-10-13	88-10-06	88-10-06	88-10-12	88-10-14	
↗				89-02-01				
↗			89-03-01		89-02-26	89-03-09	89-03-12	
↘			89-03-29	89-04-07	89-04-04	89-04-08	89-04-11	
↗		89-05-12	89-05-02	89-04-23	89-04-24	89-04-16	89-04-13	
↗	89-05-24							↗ 89-05-31
↘	89-07-16	89-07-10	89-07-06	89-06-20	89-06-23	89-07-07	89-07-09	
								↗ 89-10-10
↗	89-10-16	89-10-17	89-10-20	89-10-27	89-10-31			
↘	89-12-17	89-12-17		89-12-06	89-12-04			
↗	89-12-24	89-12-23		90-01-03	89-12-27	90-01-05	90-01-05	↗ 89-12-25
								↗ 90-03-20
↘	90-04-28	90-04-31	90-05-04	90-05-05	90-05-03	90-05-07	90-05-08	
↗	90-08-10	90-08-06	90-07-30	90-07-24	90-07-20	90-07-25	90-07-23	
								↗ 90-08-30
↘		90-12-21	90-12-20	90-12-02	90-11-30	90-11-30	90-12-02	
↘	91-02-05							
↗	91-02-17							
↗		91-03-30	91-04-10	91-04-05	91-03-28	91-04-05	91-04-10	
								↘ 91-07-01
↘	91-09-09	91-09-06	91-09-10	91-08-30	91-08-28	91-08-27	91-08-28	
↗	91-11-14	91-11-18	91-11-26			91-11-12		↘ 91-11-14
↘						91-12-03		
↘	91-12-30	91-12-30	91-12-28					↘ 91-12-30
↗				92-02-25	92-02-23			
↗						92-03-19	92-03-23	↘ 92-04-02
↗			92-04-15					
↗	92-05-10	92-04-28						
↘				92-07-07	92-07-10	92-07-12	92-07-16	
↘	92-08-03	92-07-27	92-07-25					↘ 92-07-27
↗	92-09-06	92-09-10	92-09-12					

Table 3. Coefficient Matrix A

3 months	6 months	1 year	3 years	5 years	7 years	10 years
0.693	0.381	0.019	-0.025	0.150	-0.062	-0.0936
0.141	0.769	0.016	0.115	0.031	-0.128	0.0026
0.053	0.245	0.603	0.071	0.086	-0.163	0.0515
0.021	0.201	-0.036	0.884	0.089	-0.137	0.0689
0.025	0.160	-0.037	0.134	0.790	-0.081	0.0946
0.000	0.133	-0.010	0.066	0.038	0.741	0.1001
0.036	0.086	-0.000	0.147	0.000	-0.086	0.8681

Table 4. Coefficient Matrix B

3 months	6 months	1 year	3 years	5 years	7 years	10 years
-0.045	-0.0980	-0.04783	-0.0014	-0.12574	0.08	0.0454
-0.076	-0.0035	-0.01673	-0.1413	-0.04465	0.21	-0.0043
-0.074	-0.1159	0.12864	-0.0638	-0.09166	0.24	-0.0775
-0.091	-0.0472	0.01298	-0.0684	-0.06411	0.16	-0.1213
-0.073	-0.0554	0.03210	-0.0819	-0.00041	0.12	-0.1536
-0.045	-0.0595	0.00501	-0.0520	-0.00717	0.12	-0.1645
-0.058	-0.0449	0.00049	-0.1181	0.02371	0.12	-0.1237

Table 5. Residual Standard Error

3 months	6 months	1 year	3 years	5 years	7 years	10 years
0.035	0.04	0.04	0.031	0.031	0.029	0.028

Table 6. Residual Correlation Matrix

	3 months	6 months	1 year	3 years	5 years	7 years	10 years
3 months	1.00	0.47	0.37	0.23	0.22	0.20	0.18
6 months	0.47	1.00	0.73	0.42	0.41	0.37	0.33
1 year	0.37	0.73	1.00	0.46	0.45	0.41	0.36
3 years	0.23	0.42	0.46	1.00	0.86	0.79	0.74
5 years	0.22	0.41	0.45	0.86	1.00	0.83	0.79
7 years	0.20	0.37	0.41	0.79	0.83	1.00	0.86
10 years	0.18	0.33	0.36	0.74	0.79	0.86	1.00

TABLE 7. Eigen Values of Residual Covariance Matrix

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
0.0044654	0.0016757	0.0008558	0.0004066	0.0002422	0.00013127	0.00010659

trade is biggest in the market. We could not find any good reason why all elements of the 6th column of the coefficient matrix B in Table.4 are positive and significantly high except the first. In particular, the 2nd and the 3rd elements are significant compared with other elements of the matrix. This implies that the rate of the 7 years two days before affects the 6 month or the 1 year rate positively. It is said that bond future and the 7 year swap are very much related since bonds with 7 years maturity are usually the cheapest among the nominates of the cheapest-to-deliver (bonds with 7 to 10 years maturity) for the bond future trading in the market. Note that the portfolio of long term swaps are usually hedged with bond futures. It is also said that the issuing banks of 5 years maturity bonds are limited to the three banks, the Nippon Credit Bank, the Industrial Bank of Japan, and the Long Term Credit Bank of Japan so that the 5 year swap rate is affected and tends to be lowered. But we could not find any such specific feature of the 5 years even from the behavior of residuals. These investigation of the results in relative to the specific features of the market will be a subject of the future research.

From the correlation matrix of the residuals shown in Table.6, we can observe how much correlation there is among the seven rates in one day. The fact that higher correlations for longer terms are is quite natural and understandable, because there is less speculation in longer term trading. We also note that correlation between 6 month and 1 year rate is high, but the 3 month rate is not so highly correlated to those two rates.

Another possible way of interpreting the correlation structure is to see the eigen value decomposition of the covariance matrix, that leads a principal component analysis. The eigen values are shown in Table.7 and the only first 4 eigen vectors are shown in Table.8 since the eigen values for the rest of the eigen vectors are quite small. We see from Table.7 and Table.8 that only the first two eigen vectors, that is, the first two principal components, are significant for all residuals. The others contribute only to the first three residuals, that is, Euro Yen interest rate residuals. More precisely, the residual vector $\epsilon(t)$ is decomposed as

$$\epsilon(t) = \sum_{j=1}^{7} \sqrt{\lambda_j}\, \epsilon^{(j)}(t)$$

with the eigen values $\lambda_j, j = 1, \cdots, 7$ and the corresponding orthogonal multiple time series $\epsilon^{(j)}(t), j = 1, \cdots, 7$ which have covariance matrices $u_j u_j^T, j = 1, \cdots, 7$, respectively. They are orthogonal to each other. Thus, as is seen from Table.8 the

Table 8. The First Four Eigen Vectors of the Covariance Matrix of Residuals

	u_1	u_2	u_3	u_4
3 months	0.2599885	-0.4014906	0.85685183	0.192176906
6 months	0.4779732	-0.4539444	-0.19496352	-0.726141302
1 year	0.4773444	-0.3520577	-0.45793555	0.658369742
3 years	0.3734923	0.3407595	0.05364768	0.008425985
5 years	0.3765874	0.3663068	0.05499962	-0.003553742
7 years	0.3264048	0.3614640	0.07069049	-0.018223919
10 years	0.2970443	0.3575489	0.08478813	-0.043822272

Table 9. Estimated Quantiles

	3 months	6 months	1 year	3 years	5 years	7 years	10 years
0.93	3.288799	3.686934	3.761160	2.997981	2.993087	3.314082	3.114515
0.95	3.767082	4.194803	4.316260	3.625486	3.672817	4.027951	3.812759
0.97	4.584951	4.969545	5.409281	4.506268	4.734835	5.101522	4.861329

first $\epsilon^{(1)}(t)$ can be considered a common factor which almost uniformly contributes to all the seven residuals and the second $\epsilon^{(2)}(t)$ can be considered as a factor which contributes in an opposite way to the last four swap rate series, because the first three elements of the 2nd eigen vector are negative and the others are all positive. In short, we can say that the first factor is a common factor, the second factor is a factor discriminating the Euro Yen and the LIBOR swap rates, and the rest of factors are small factors which contribute only to the Euro Yen rates.

As is seen in Figure.10, the distribution of residuals is quite heavily tailed, which can be approximated by a stable distribution with the exponent around $\alpha = 1.3$. Table.9 shows the estimates $\hat{z}_f$ of the f quantile of the standardized symmetric stable distribution with exponent α,

$$\hat{z}_f = (.827)\frac{\hat{x}_f - \hat{x}_{1-f}}{\hat{x}_{.72} - \hat{x}_{.28}}$$

for $f = 0.93$, $f = 0.95$ and $f = 0.97$ (Fama and Roll, 1971), where $\hat{x}_f$ is the f quantile of the residuals. Simple estimates of α can be obtained by looking into Table.1 in Fama and Roll(1968). The estimates varies according to the value of f but fall in the range from 1.1 to 1.4 for all the seven series. These estimates of α

Table 10. Estimates of the Exponent α of Stable Distribution

	3 months	6 months	1 year	3 years	5 years	7 years	10 years
0.93	1.2	1.1	1.1	1.3	1.3	1.2	1.2
0.95	1.3	1.2	1.2	1.3	1.3	1.2	1.3
0.97	1.4	1.4	1.3	1.4	1.4	1.4	1.4

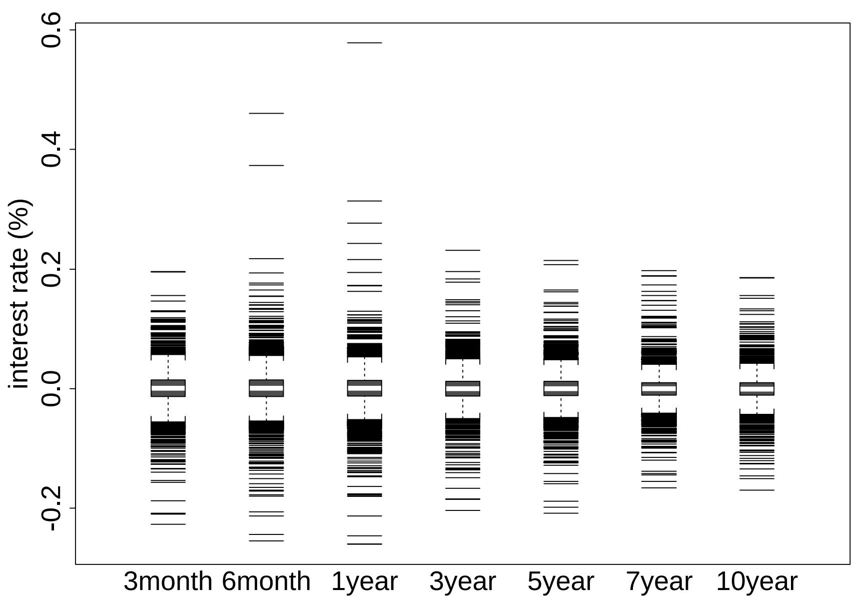

FIGURE 10. Boxplots of residuals

Table 11. Variability of Each Component of the Decomposition

	3 months	6 months	1 year	3 years	5 years	7 years	10 years
long term trend	0.00716	0.00756	0.00827	0.00797	0.0078	0.00688	0.00709
short term trend	0.00518	0.00512	0.00499	0.00797	0.0078	0.00688	0.00709
irregular	0.07485	0.07611	0.07827	0.06976	0.0680	0.06345	0.05705

are shown in Table.10. Since no optimal f is known, we may satisfy ourselves by saying that the α is around 1.3. However, it is interesting to compare this α with that for return rate of stock prices, $\alpha = 1.7$ (Mandelbrot, 1963). The residuals here have much heavier tailed distribution than that of the return rate of stock prices.

5. Prediction

Using the fitted MAR(2) model in Section 4.3, we can predict interest rates on a daily basis. For example, a practical one day ahead prediction of seven series is given at once by

$$\hat{Z}(t+1) = L(t) + S(t) + A\,I(t) + B\,I(t-1).$$

This is due to the fact that the variability of the irregular component dominates those of the long term trend and the short term trend as is seen in Table.11. Here, the variabilities for the long and short terms are the standard deviations of the lag 1 differences in regarding that those are random fluctuations, and the variabilities for the irregular is exactly the standard deviations of the stationary process $I(t)$. Since the $Z(t+1)$ can be represented as

$$Z(t+1) = L(t+1) + S(t+1) + A\,I(t) + B\,I(t-1) + \epsilon(t+1)$$

the prediction error becomes

$$Z(t+1) - \hat{Z}(t+1) = \{L(t+1) - L(t)\} + \{S(t+1) - S(t)\} + \epsilon(t+1)$$

and the first two terms on the right hand side of the above equation are negligible compared with the last term $\epsilon(t+1)$. The standard error of $\epsilon(t+1)$ is around 0.03 to 0.04 as is seen in Table.5, so that the prediction error is reduced to almost half by predicting by $\hat{Z}(t+1)$ based on the MAR(2) model, if it is compared with that by a simple prediction based only on the long term and short term trends,

$$\tilde{Z}(t+1) = L(t) + S(t).$$

We note that we neglected the estimation error of coefficients A and B in the discussion above. Such errors are negligible in order of magnitude, because we estimated those based on long enough observation for 2115 days.

On the other hand, if we are only interested in long term forecasting, then a forecasting based on the long and short term trends would be more stable than that based on the whole data, because the variability of irregular component is much larger than that of the long and short term trends.

6. Applications in Financial Engineering

Interest Rate Options Pricing

The series of daily increment of the interest rates and the swap rates can be approximately expressed by a multivariate autoregressive model of order 2 on the basis of the analysis given in this paper. On the other hand, a commonly used Ornstein-Uhlenbeck process is a stochastic model of continuous time series and an autoregressive model of order 1 is frequently employed for modeling such data as economic indices by an analogy to the Ornstein-Uhlenbeck process. Our analysis shows that autoregressive model of order 1 does not work well even for describing the irregular series, because the lag 1 and the lag 2 coefficient matrix A and B are both significant, which are though quite free from non-stochastic component. This can be understood if we recall the fact that sampling from continuous Ornstein-Uhlenbeck process does not necessarily follow an autoregressive process of order 1, rather follows an autoregressive moving average process. Our MAR(2) might be interpreted as an approximation to such an autoregressive moving average process.

Furthermore, the principal component analysis applied to the residuals of the MAR(2) in Section 4 provides that they have at least two main factors. This may relate to the topic of the pricing theory of the derivatives with multi-factor stochastic interest rates (for example, see Miura and Kishino, 1995 and Duffie, 1992).

Risk Management

Tomorrows interest rates and swap rates can be predicted by the formula given in Section 5, with the error described. Those errors are the stochastic parts of the daily move of the interest rates. The variance and covariance of the residuals of MAR(2) or the prediction error can be used for the measurement (especially the estimation of the volatility) of the daily change of the value of the financial assets. For example, they can be used for computing the Value at Risk (see Morgan Guaranty Trust Company, 1994).

7. Technical details

In this last section, we will explain some technical details of our data analysis for those who are interested in re-analyzing our data or analyzing similar or updated data. However, the following explanation does not perfectly describe our real process of analyses because the real process is usually a series of trials and errors, which can not, in general, be well organized beforehand.

7.1. Data Cleaning. An ASCII file of the original data was obtained by the courtesy of the Long Term Credit Bank of Japan. This raw data file is now available for public from `stat.math.keio.ac.jp:/usr/pub/statlib/s.jpn` by anonymous

FTP. After reading the file as a list in S by a function `scan`, we interpolated by S function `approx` values for Sundays and holidays as well as for missing days. We also adjusted the number of days in a year into 365 days by omitting values for the 29 th of February in leap years. The starting date is adjusted to the 1st of December, 1986 since in some of the series the data are not available prior to this date. The official discount rate was not included in the file so that we picked up it from news papers and organized as an S data set, too. Then we organized the seven interest rates as a list `yen.int` of time series objects together with the official discount rate, in which each data has a time series attribute. The S list `yen.int` is also available as an S dump file `yen_int.s` together with the `yen_int.d` in a compressed tar file `yen_int.tar.Z` by anonymous FTP as mentioned above.

7.2. Analysis. The decomposition of each series into three components, "long term trend", "short term trend" and "irregular" is quite simply obtained by using the following S function.

```
"decomp"<-
function(original, f1 = 365/length(original), f2 = f1/12)
{
#
# original: original time series
# f1: fraction of a yearly span
# f2: fraction of a monthly span
#
lowess1 <- lowess(original, f = f1)
long.trend <- lowess1$y
tsp(long.trend) <- tsp(original)
long.irregular <- original - long.trend
lowess2 <- lowess(long.irregular, f = f2)
short.trend <- lowess2$y
tsp(short.trend) <- tsp(original)
irregular <- long.irregular - short.trend
decomp <- list(long.trend = long.trend, short.trend = short.trend,
irregular = irregular)
class(decomp) <- "decomp"
return(decomp)
}
```

Fitting an MAR model to the vector time series of seven irregular series is done by Splus (which is a commercial version of S) function `ar`. This function selects the order of the autoregression so as to minimize AIC. We tried other larger orders like 3 or 4 than the 2 which is selected by this function, but we could not find any significant difference.

The authors are indebted to Mr. Hirokazu Sugimoto at the Long Term Credit Bank of Japan for his conversation with us on the behavior of interest rates.

A part of the results in this paper is previously reported in Shibata, Miura and Uchida (1993).

References

Becker R.A., Chambers J.M. and Wilks A.R. (1988) *The New S Language*, Wadsworth, California.

Beveridge, S. and Nelson C. (1981) A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle, *Journal of Monetary Economics*, 7, 131-174.

Cleveland, W.S. (1979) Robusts locally weighted regression and smoothing scatterplot, *J. Am. Statist. Assoc.*, 74, 829-836.

Cleveland, W.S., Devlin, S.J., Schapira, D.R. and Terpenning, I.J. (1981) The SABL seasonal and calendar adjustment package, Computing Information Library, Bell Laboratories.

Cleveland, W. S. and Devlin, S. J. (1982) Calendar effects in monthly time series: modeling and adjustment, *J. Am. Statist. Assoc.*, 77, 520-528.

Cleveland, W. S. and Devlin, S. J. (1988) Locally weighed regression: an approach to regression analysis by local fitting, *J. Am. Statist. Assoc.*, 83, 596-610.

Cochrane, J. (1988) How big is the random walk in GNP?, *J. Political Economy*, 96, 893-920.

Duffie, D. (1992) *Dynamic Asset Pricing Theory*, Chapter 7, Princeton University Press, Princeton.

Enders, W, (1995) *Applied Econometric Time Series*, John Wiley and Sons.

Fama, E. and Roll, R. (1968) Some properties of symmetric stable distributions, *J. Am. Statist. Assoc.*, 63, 817-836.

Fama, E. and Roll, R. (1971) Parameter estimates for symmetric stable distributions, *J. Am. Statist. Assoc.*, 66, 331-338.

Morgan Guaranty Trust Company (1994) *Risk Metrics. Technical Document*, 2nd ed., J.P. Morgan, New York.

Mandelbrot, B. (1963) The variation of certain speculative prices, *J. Business*, 36, 394-419.

Miura, R. and Kishino H. (1995) Pricing of bonds and their derivatives with multi-factor stochastic interest rates: A note, in *Nonlinear and Convex Analysis in Economic Theory*, eds. T. Maruyama and W. Takahashi, Lecture Notes in Economics and Mathematical Systems, 419, 215-229, Springer.

Nelson, C. and Plosser, C. (1982) Trends and random walks in macroeconomic time series, *J. Monetary Economics*, 10, 129-162.

Quah, D. (1992) The relative importance of permanent and transitory components: identification and some theoretical bounds, *Econometrica*, 60, 119-143.

Shibata, R., Miura, R. and Uchida, K. (1993) Decomposition of interest rate series through local regression, *Proceedings of the Annual Meeting of the Japan Statistical Association.*

Department of Mathematics, Keio University, 3-14-1 Hiyoshi Kohoku, Yokohama, 223, Japan

Department of Commerce, Hitotsubashi University, 2-1 Naka Kunitachi, Tokyo, 186, Japan