

# 1 Statistics of Market Rates

The bulk of OTC derivatives consist of *interest rate swaps* and *cross-currency swaps*. A cross currency swap is similar to a (single currency) interest rate swap, except that the payments made by the two parties are in different currencies. Either side may involve fixed interest payments or interest payments tied to a floating rate. There is an agreed “notional principal” amount in each currency, and typically the swap starts and finishes with the actual exchange of these sums. The notional principals are usually chosen to be equal in value at the inception of the swap, but changes in foreign exchange (FX) rates may make them very unequal at its maturity. The “mark to market” value of the swap therefore depends on FX rates. If either side of the swap carries a fixed rate, the value also depends on the yield curve in the corresponding currency. A floating rate side is relatively insensitive to yield curve changes.

One simple kind of cross-currency swap is the *FX forward contract*. This is an agreement to exchange specified sums in two currencies on a specified future date. It therefore involves only a single cash flow on each side, at maturity. Because its value depends on the FX rate and the zero-coupon yields for the specified maturity in both currencies, the value for a long-term forward is very volatile. To avoid the yield-curve sensitivity, long term hedging of FX risk is usually effected through a cross-currency swap with floating rate payments on both sides—a *basis swap*. Although the cash flows are more complicated than those for a forward, the elimination of yield-curve sensitivity usually more than compensates for the additional complication.

These two basic types of swap may account for 90% of the volume of a swaps market participant. The balance would typically be options of one sort or another: FX options, that is puts and calls on FX rates; interest rate caps and floors; and *swaptions*, that is options to enter swaps. The values of options depend on volatilities of underlying rates as well as on the rates themselves.

## 1.1 Foreign Exchange Rates

Figure 1 shows the JPY/USD exchange rate series from January, 1995, to the start of 1998. The rate is expressed as the price of \$1.00 in Yen, and measures the strength of the US dollar relative to the yen; the low around April, 1995, reflected a rapid slide in the dollar that started in March, followed by a

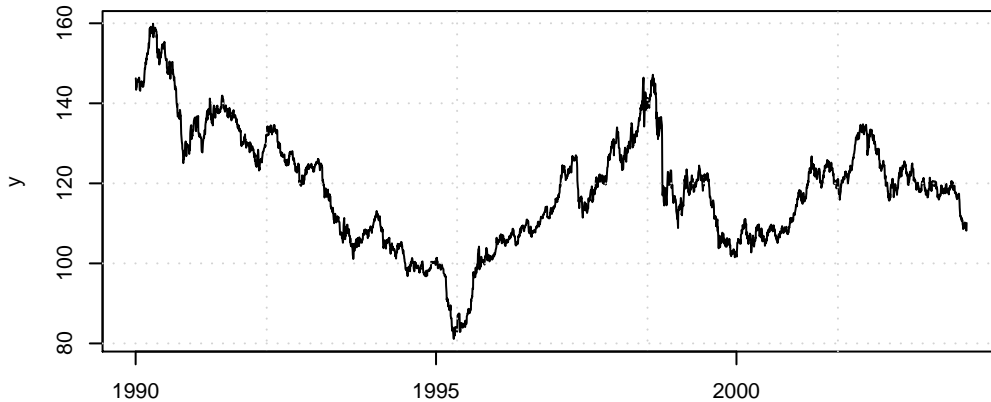


Figure 1: Dollar/yen exchange rate.

matching recovery.

It seems reasonable to study the day-to-day changes in the rate, and the logarithm of the ratio of one day's rate to the previous rate, shown in Figure 2, suggests itself. Figure 3 shows a *qq*-plot of these differences against the normal distribution (that is, a graph of the quantiles of the empirical distribution of logarithmic differences *versus* matching quantiles of the normal or Gaussian distribution). The distribution is substantially longer tailed than the Gaussian, as indicated by the distinct curvature.

The next question concerns the correlation structure of the series. Figure 4 shows the correlogram of the series of logarithmic differences, with lines indicating the limits for testing at the 5% level the hypothesis that the autocorrelation is zero. Only one correlation (out of 28) appears to be definitely significant at the 5% level, in line with what we would expect under the null hypothesis. Interestingly, the first correlation, at lag 1, is marginally significant and negative. This suggests that the market overreacts to an event on a given day, and retreats slightly on the following day. However, one would not want to over-interpret the graph.

The original graph (Figure 1) suggests that in the long term the series

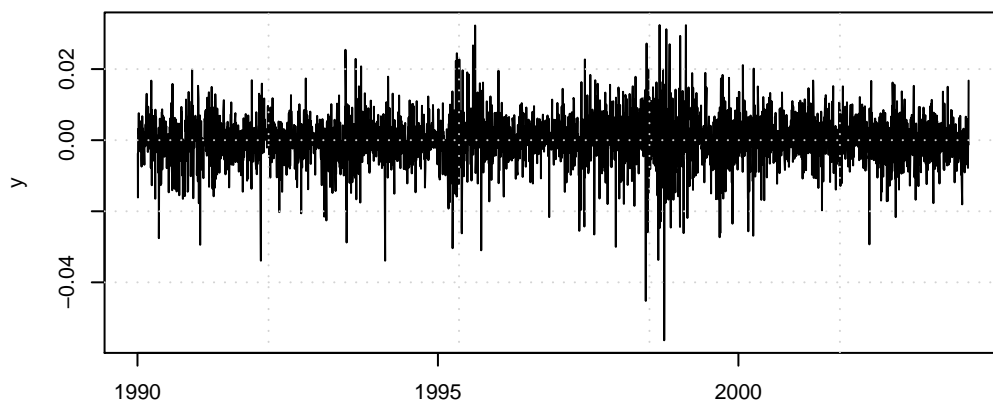


Figure 2: Logarithmic returns, dollar/yen exchange rate.

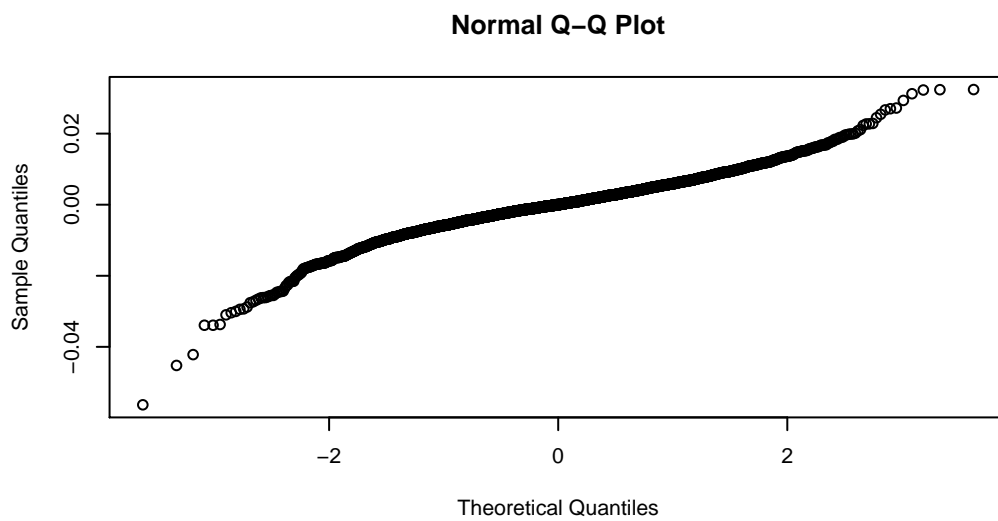


Figure 3: Quantile-quantile plot of logarithmic differences of dollar/yen exchange rate.

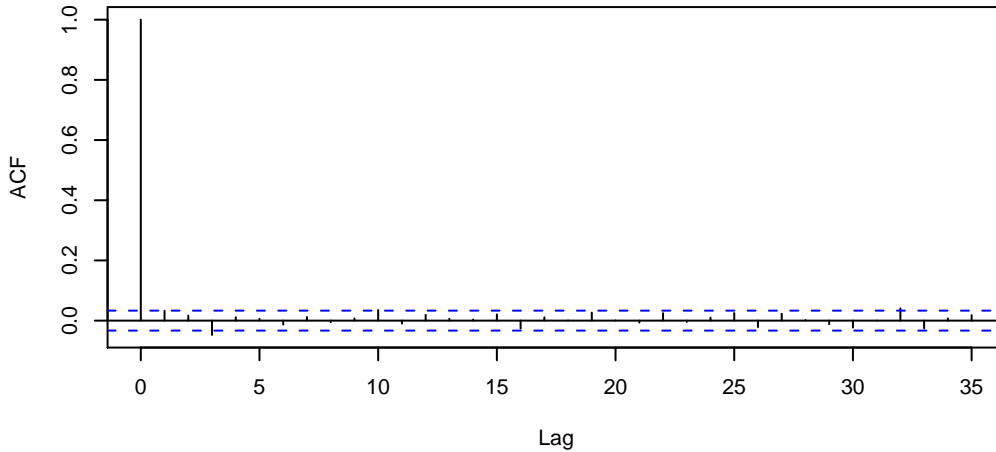


Figure 4: Correlogram of logarithmic differences of dollar/yen exchange rate.

may be subject to trends that are sustained for some months. If this were true it would lead to small *positive* autocorrelations for several lags, and there is in fact a preponderance of positive (though insignificant) correlations in the graph. A good way to explore such effects is through the *power spectrum*. Figure 5 shows an estimate of the power spectrum of the series of logarithmic differences. The spectrum units are *decibels*, or  $10 \times \log_{10}(\text{spectrum})$ ; 1 decibel corresponds to a factor of 1.26 in the power spectrum. The vertical line towards the right hand side of the graph is a (pointwise) 95% confidence interval for the spectral density, and indicates that the data are consistent with a flat spectrum, or in other words that the logarithmic differences appear to be white noise (having no serial correlation).

Sustained trends on a time-scale of months or years would have been expected to introduce a peak in the spectrum at low frequencies. In the graph, frequencies are measured on a scale of cycles per day, so a monthly cycle would occur at the frequency 1 cycle per month or roughly 0.05 cycles per day. A trend sustained for one month would not correspond to any exact cycle, but would be broken down across frequencies from zero to roughly one half this, or 0.025 cycles per day. The absence of any such peak (at least,

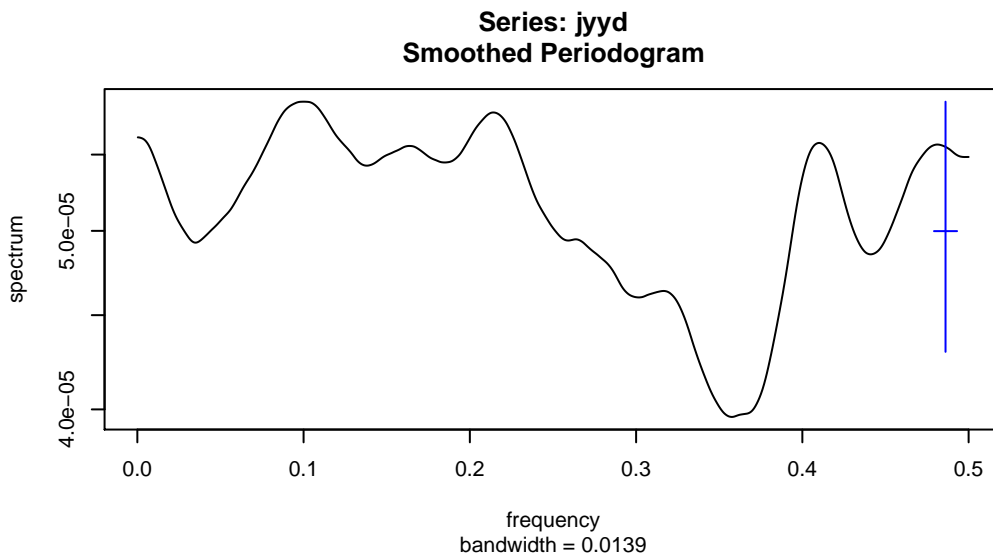


Figure 5: Power spectrum of logarithmic differences of dollar/yen exchange rate.

one that reaches higher than peaks at much higher frequencies) means that the trends in the original figure are no stronger than one would expect on the basis of a random walk: accumulated white noise.

## 1.2 Conditional Heteroscedasticity

We have seen that the JPY/USD exchange rate exhibits nearly uncorrelated day-to-day logarithmic changes. However, this does not mean that the change from one day to the next is *independent* of the past. In fact, the *magnitudes* of the changes show a certain degree of dependence. Figure 6 shows the correlogram of the absolute values of the changes, and Figure 7 shows an estimate of the corresponding power spectrum. These graphs show that there is some degree of persistence in the volatility of the day-to-day changes. For instance, the correlation between a given absolute change and the average of the preceding 10 days' absolute changes is around 0.2. Furthermore, if we examine the distribution of the ratio of a given change to the average of the preceding 10 days' absolute changes, the *qq*-plot is as shown in Figure 8. This is somewhat straighter than the plot for the original data, except perhaps

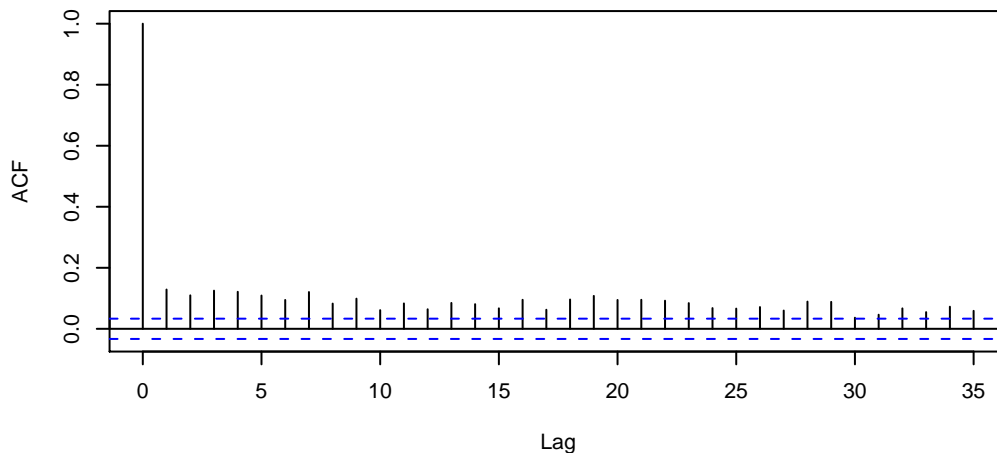


Figure 6: Correlogram of absolute values of the logarithmic differences of dollar/yen exchange rate.

for a small number of points at each end, and suggests that it may be more reasonable to view the changes as being *conditionally* normally distributed, given the past values, than as *unconditionally* normally distributed. (Note that the partial straightening of the *qq*-plot cannot be an artefact: dividing a variable by another unrelated variable can only *increase* the heaviness of the tails.) Figure 9 shows a *qq*-plot of the same ratios against the *t*-distribution with 5 degrees of freedom, and suggests that this distribution might provide a better model.

This characteristic is called *conditional heteroscedasticity*. We might write

$$e_t | e_{t-1}, e_{t-2}, \dots \sim N(0, \sigma_t^2)$$

where  $\sigma_t$  is some function of  $e_{t-1}, e_{t-2}, \dots$ , or more constructively

$$e_t = \sigma_t z_t$$

where  $z_t \sim N(0, 1)$  and the  $z$ 's are independent of each other. Again, a longer tailed distribution than the normal would provide a closer model.

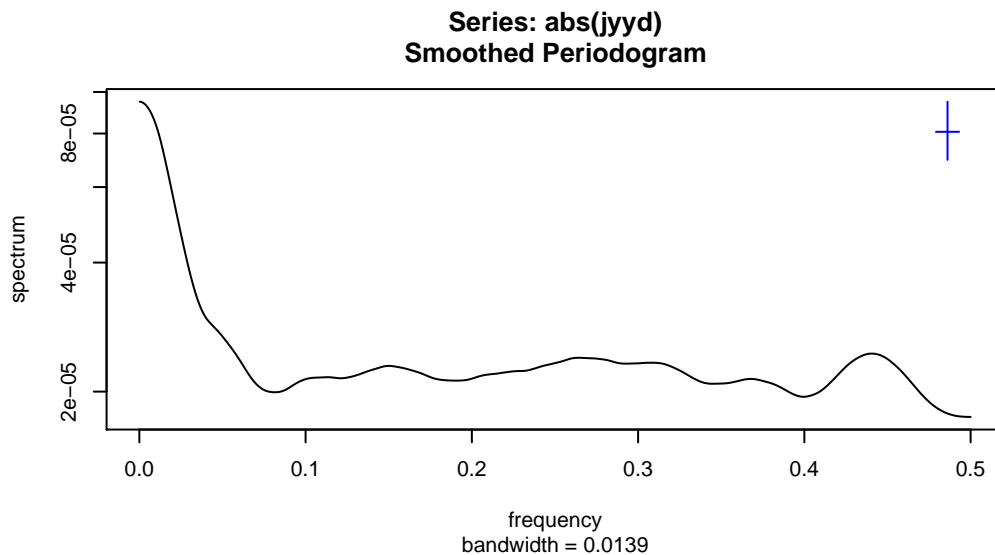


Figure 7: Power spectrum of absolute values of the logarithmic differences of dollar/yen exchange rate.

### 1.3 ARCH and GARCH Models

The Auto-Regressive Conditionally Heteroscedastic (ARCH) model is related to this structure. It consists of an autoregressive structure for some observed time series  $\{y_t\}$ , in which the errors  $\{e_t\}$  have the above structure:

$$y_t = ay_{t-1} + e_t,$$

for instance. In the original ARCH model, the function  $\sigma_t^2$  was just a linear compound of recent  $e^2$ 's:

$$\sigma_t^2 = \omega + \sum_{r=1}^q \alpha_r e_{t-r}^2$$

However, it was quickly recognized that this could be interpreted as having  $\sigma_t^2$  depend in a Moving Average (MA) fashion on the  $e^2$ 's. The Generalized ARCH (GARCH) model extends this dependence to something similar to ARMA:

$$\sigma_t^2 = \omega + \sum_{r=1}^p \beta_r \sigma_{t-r}^2 + \sum_{r=1}^q \alpha_r e_{t-r}^2.$$

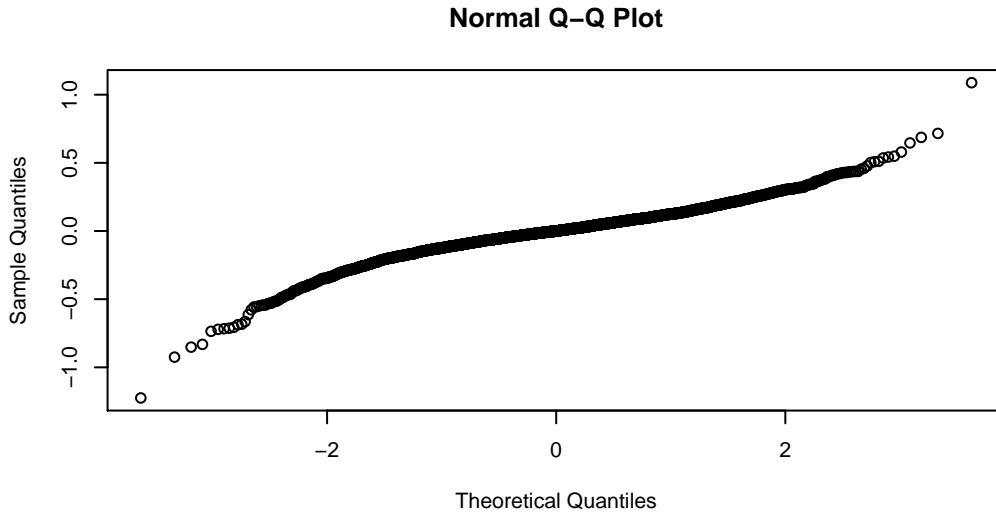


Figure 8: Quantile-quantile plot of ratio of logarithmic differences of dollar/yen exchange rate to the average of the preceding 10 absolute differences.

Of course, the underlying  $z$ 's need not be normally distributed. The  $qq$ -plot shown above shows some remaining heavy-tailedness; this might be attributed to inadequate modelling of the conditional variance, or could indicate that there is intrinsically nonGaussian behavior (including a hint of skewness).

Fitting a GARCH model requires identifying good values for  $p$  and  $q$  and estimating the  $\alpha$ 's and  $\beta$ 's, in addition to the modelling of the autoregressive part (the equation relating  $y$ 's to  $e$ 's), which could itself be in the general ARMA form. Identification and estimation are nontrivial exercises; modules for carrying out these steps are available for R, Splus, and SAS.

Fitting the GARCH(1, 1) model using the R function `garch` from the `tseries` library gives the output shown in Figure 10, which shows that

$$\hat{\omega} = 7.995 \times 10^{-7}, \hat{\alpha}_1 = 0.03932, \hat{\beta}_1 = 0.9442.$$

A GARCH(1, 1) model is stationary iff  $\alpha_1 + \beta_1 < 1$ , so this fitted model is close to non-stationarity. The *unconditional* variance is

$$\omega/[1 - (\alpha_1 + \beta_1)],$$

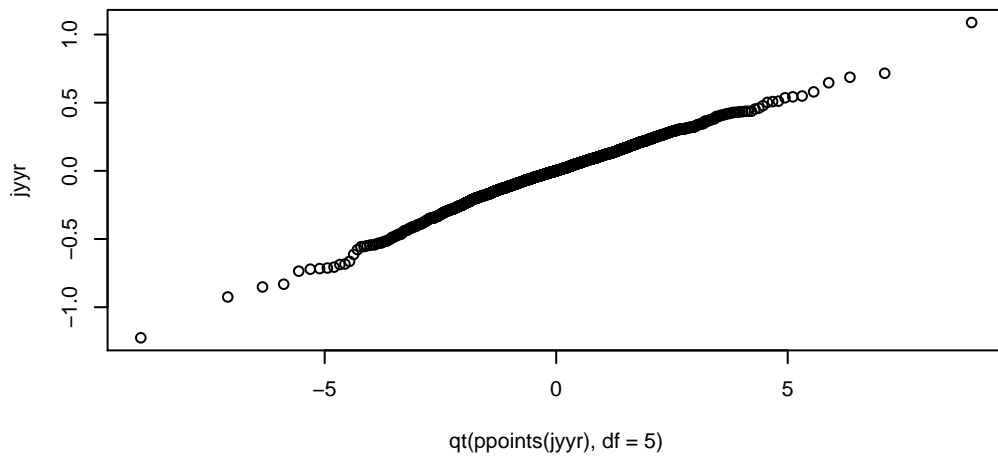


Figure 9: Quantile-quantile plot of ratio of logarithmic differences of dollar/yen exchange rate to the average of the preceding 10 absolute differences against the  $t$ -distribution with 5 degrees of freedom.

```

Call:
garch(x = jyyd)

Model:
GARCH(1,1)

Residuals:
      Min       1Q   Median       3Q      Max
-6.0105 -0.5487  0.0000  0.5904  4.1413

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
a0 7.995e-07  1.221e-07   6.549 5.8e-11 ***
a1 3.932e-02  3.760e-03  10.458 < 2e-16 ***
b1 9.442e-01  5.388e-03  175.263 < 2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Diagnostic Tests:
Jarque Bera Test

data: Residuals
X-squared = 1086.034, df = 2, p-value < 2.2e-16

Box-Ljung test

data: Squared.Residuals
X-squared = 1.3951, df = 1, p-value = 0.2375

```

Figure 10: GARCH(1, 1) output.

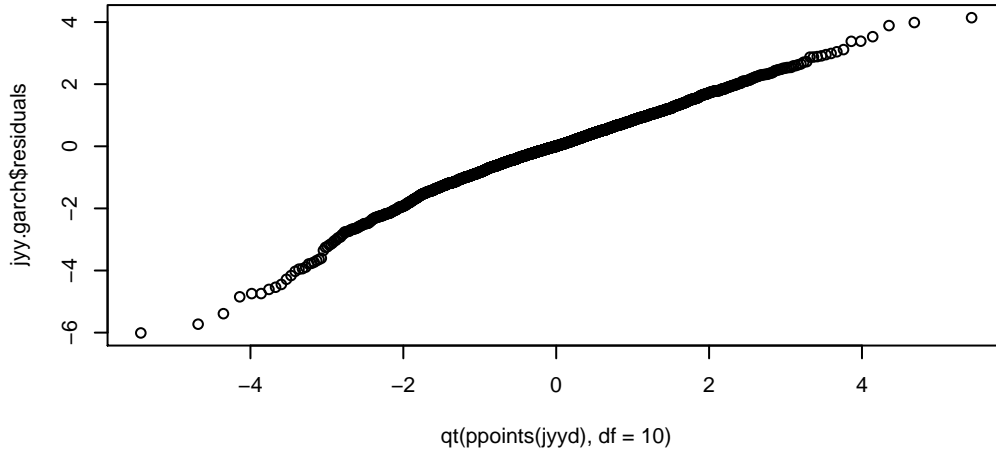


Figure 11: Quantile-quantile plot of GARCH residuals against the  $t$ -distribution with 10 degrees of freedom.

which is estimated here as  $4.866 \times 10^{-5}$ , for a daily volatility of 0.006975 and an annualized volatility of 0.1103 (the daily volatility multiplied by  $\sqrt{250}$ ). The Jarque Bera test, a test for departures from the normal distribution, is not unexpectedly highly significant. The Box-Ljung test gives no evidence that the GARCH(1, 1) formulation is inadequate.

The  $qq$ -plot of the GARCH residuals against the  $t$ -distribution with 10 degrees of freedom is shown in Figure 11. The asymmetry of the distribution is now clear. The upper tail is fitted well, but the lower tail is longer than that of the matching  $t$ -distribution.

## 1.4 Yield Curves

Figures 12 and 13 show two views of the US yield curve from the end of 1996 to early August, 2003. The set of maturities is: 3mo; 6mo; 1yr; 2yr; 3yr; 5yr; 7yr; 10yr; 15yr; 30yr.

The yields for maturities of a year or less are LIBOR rates. Data for longer maturities are “swaps curve” data, a version of the yield curve that

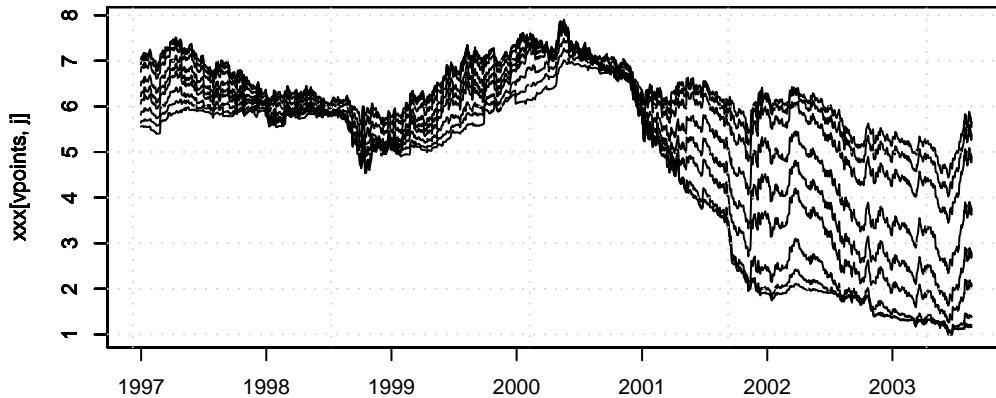


Figure 12: US yield curve.

is adapted for use in an interest rate swap (IRS). The entry on a given date and for a given maturity is the interest rate for the *fixed* side of an IRS, as set by the market on that date. That is, an entity wishing to begin an IRS, fixed *versus* LIBOR flat, would have been quoted this rate for the fixed side, if the swap was to have been entered “at the market”.

The rates are averages of quotes from a number of dealers. These are “mid-market” rates, that is the average of the “bid” (lower) and “ask” (higher) rates. The difference between the bid and ask rates is the *bid-ask spread*, typically around 2 *basis points*, or 0.02%, for USD swaps.

Because the fixed side of an IRS has periodic payments, the swaps curve is a *current coupon* yield curve rather than a *zero coupon* curve. However, as we have noted elsewhere, one may be obtained from the other without difficulty. The swaps curve is typically 20-30 basis points higher than the Treasury yield curve, reflecting the presence of a small amount of risk that is absent in Treasury securities.

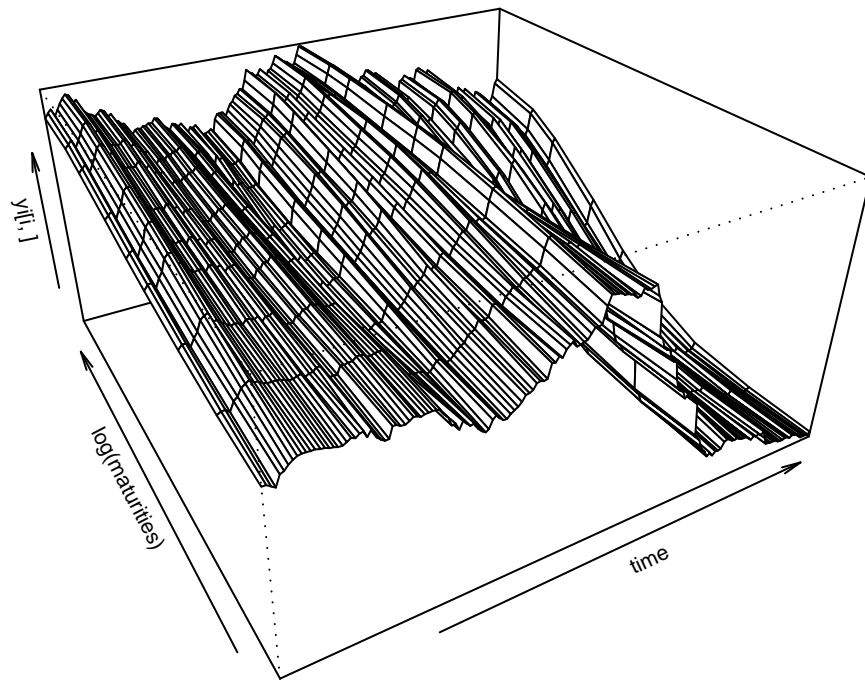


Figure 13: Perspective view of US yield curve.

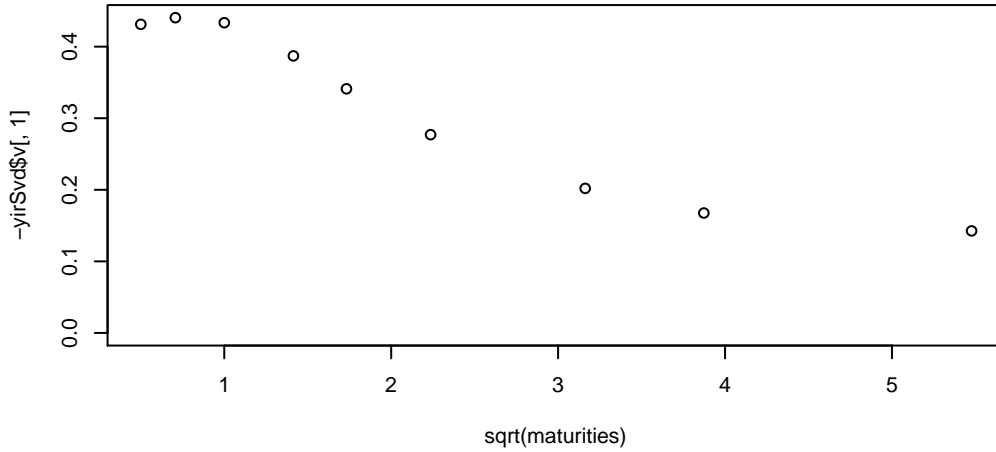


Figure 14: Loadings for first PC of US yields.

## 1.5 Statistical treatment

A yield curve may be thought of as a multivariate random variable. One obvious summary of data such as these is through principal components analysis (PCA). Figures 14 and 15 show the loadings for the first two components, which account for 96.4% and 3.4% of the total variance, respectively. They are graphed against the square root of maturity, for graphical rather than theoretical reasons.

The loadings for the first component are all positive, and therefore represent *shifts*: a tendency for yields to move in tandem across all maturities. The loadings are all of similar magnitude for maturities of a year or more, and to this extent represent *parallel* shifts. The smaller loadings at shorter maturities imply that these rates are less strongly affected by this dominant mode of variation. The second set of loadings is negative at the short end of the curve, and positive at the long end, and represents *tilts*: a tendency for the slope of the yield curve to vary. Note that these tilts would be *linear* only on a time scale on which the 10-year and 30-year rates are very close together.

It may be more useful to look at the yield curves in terms of day-to-

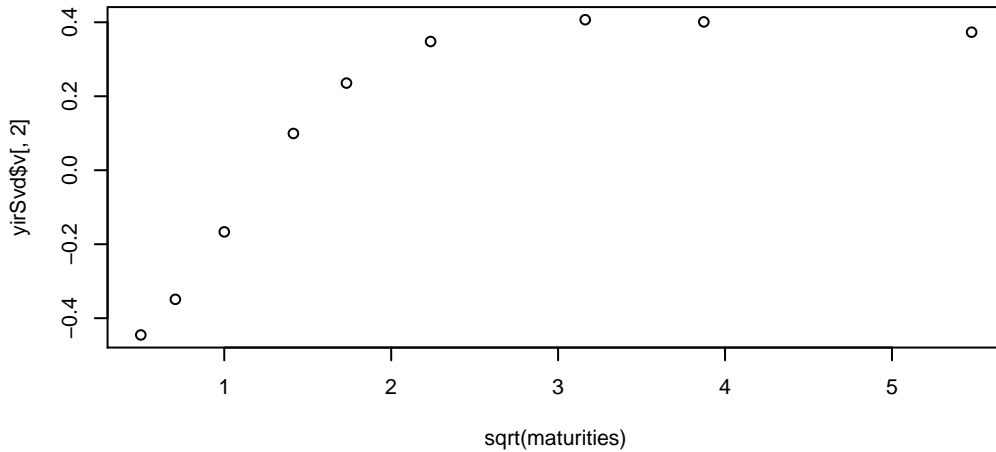


Figure 15: Loadings for second PC of US yields.

day changes, as it generally is for foreign exchange rate data. The first two components account for 85.1% and 9.5% of variance in this case (the third accounts for 3.1%). The loadings, shown in Figures 16 and 17, are still broadly interpretable as a shift factor and a tilt factor, but evidently the behavior is not as simple as when studying the levels of yields themselves.

Note in particular that there appears to be a break in behavior between maturities of one year and less *versus* the longer maturities. This reflects the different markets that determine LIBOR rates and swap rates.

In both pairs of loadings, it would presumably be possible to find linear combinations that more closely represent parallel shifts and linear tilts, respectively. These would jointly capture the same total fraction of the overall variance, but because of a loss of orthogonality it would no longer be possible to distribute this fraction unambiguously between them. This minor inconvenience might be a worthwhile price to pay for a more conventional interpretation of the decomposition.

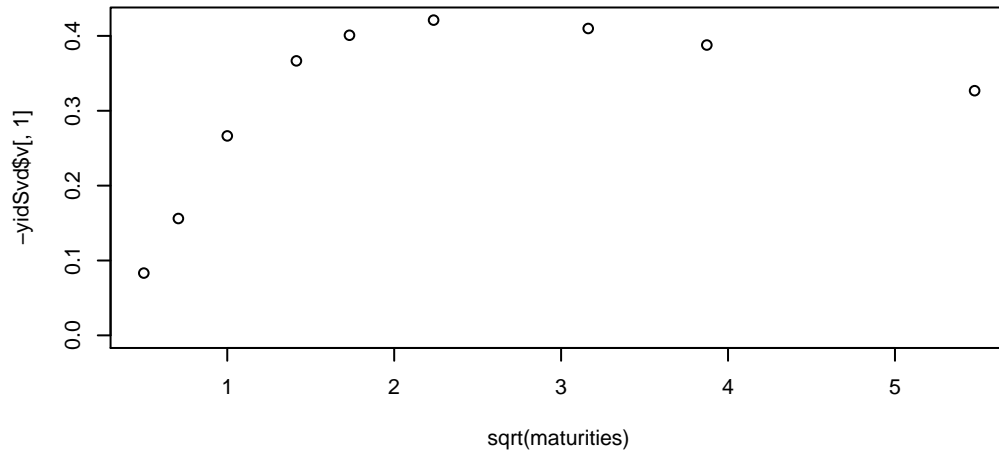


Figure 16: Loadings for first PC of day-to-day changes in US yields.

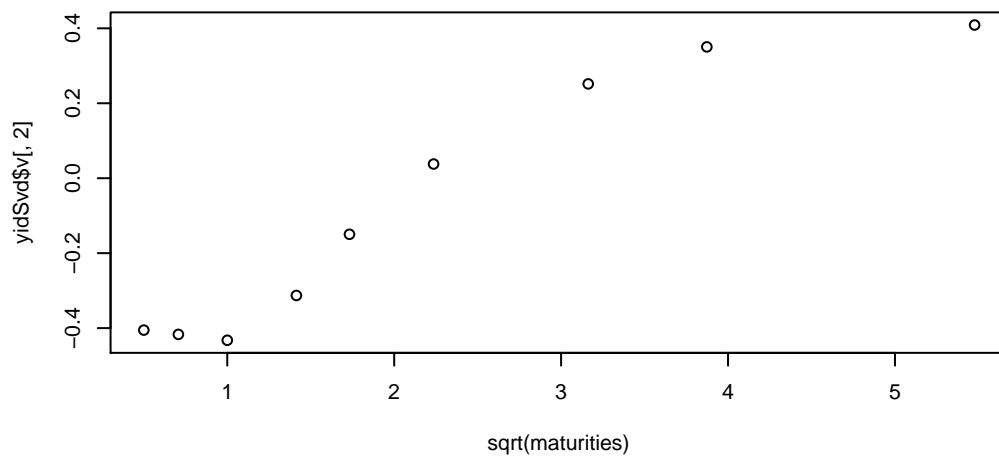


Figure 17: Loadings for second PC of day-to-day changes in US yields.

## 1.6 Looking forward: modeling and simulation

The analyses of the previous section suggest that a statistical model for yield curves might be based on a low-dimensional representation derived from PCA or a similar method. Each component is associated with a time series of scores, which would need to be modeled in the same general way as was discussed earlier for FX rates.

There are, however, constraints that should be imposed. In the case of FX rates, the only constraint is that they should be positive, and this is easily imposed by modeling the log rate. Interest rates are also intrinsically positive, again suggesting modeling on the log scale. Swaps payments may of course be negative, for instance if they are set according to a floating rate less a spread, unless there is a floor (express or implied) at zero. Low rates in Japan have made this apparently academic possibility a real one.

Interest rates however also satisfy some other constraints. The zero-coupon yield curve is related to the discount curve by  $D(t) = e^{-tr(t)}$ , and it is natural for  $D(t)$  to be required to be monotonically decreasing. One way to represent this is in terms of the *implied forward* rates. For  $0 < t < u$ , write

$$r(t, u) = \frac{\log D(u) - \log D(t)}{u - t}$$

whence

$$D(u) = D(t)e^{-(u-t)r(t,u)}.$$

The quantity  $r(t, u)$  is referred to as the forward rate at time  $t$  for a loan maturing at time  $u$ . A deposit that earned the rate  $r(t)$  until time  $t$  and was then deposited for the additional time  $u - t$  at a rate of  $r(t, u)$  would accumulate to the same value as if it were deposited until time  $u$  at the current rate for such deposits,  $r(u)$ . This is the sense in which the rate is *implied* for forward deposits. It is then clear that *the discount curve is monotonically decreasing if and only if all implied forward rates are positive*.

It is also clear that any of the sets of data:

- $D(t_1), D(t_2), \dots, D(t_n)$ ;
- $r(t_1), r(t_2), \dots, r(t_n)$ ;
- $r(t_1), r(t_1, t_2), \dots, r(t_{n-1}, t_n)$ ;

may be obtained from any other, and amount to alternative representations of the same yield curve. Since the last satisfy the simplest constraints, modeling

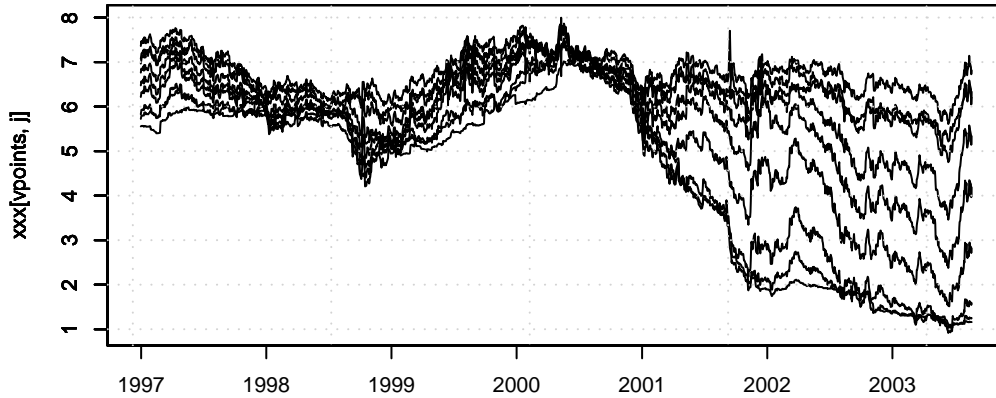


Figure 18: US implied forwards.

and simulating the logarithms of the implied forward rates may be the most direct way of incorporating the monotonicity property.

Figure 18 shows the implied forward rates corresponding to the rates shown in Figure 12. While they are different from the original yields, there is a broad similarity, and PCA also shows broadly similar results. The loadings for the first component are shown in Figure 19. As before, the loadings are all positive, although they range in magnitude. Figure 20 shows the corresponding time series component. The graph reveals two features that may require modelling:

- in some time intervals, the changes appear to be centered around non-zero values; if the serial correlations are significant, a conditional mean (ARMA) model may be needed;
- the volatility is clearly not constant, and a conditional variance model is needed, perhaps for the residuals from an ARMA conditional mean model.

However, the conditional variance structure does not necessarily mean that we should fit a CH model. The increase in volatility in late 2001 coincides

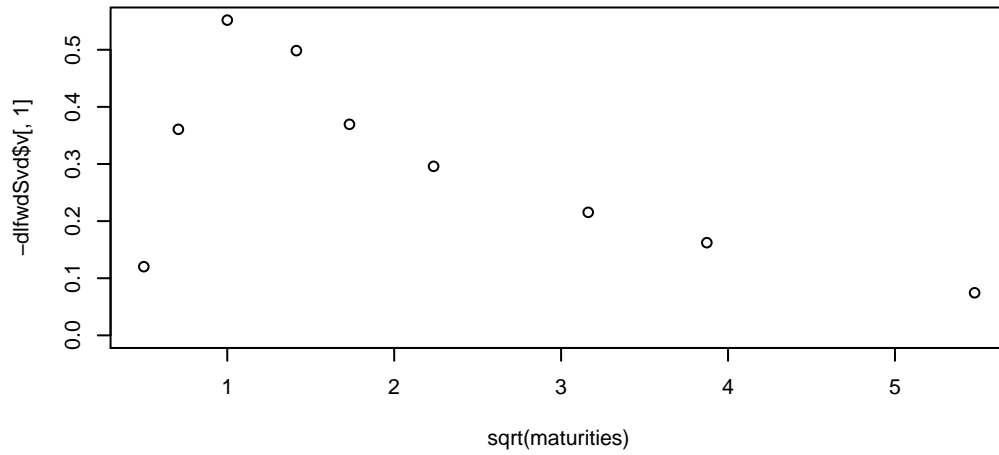


Figure 19: Loadings for first PC of day-to-day changes in log forward rates.

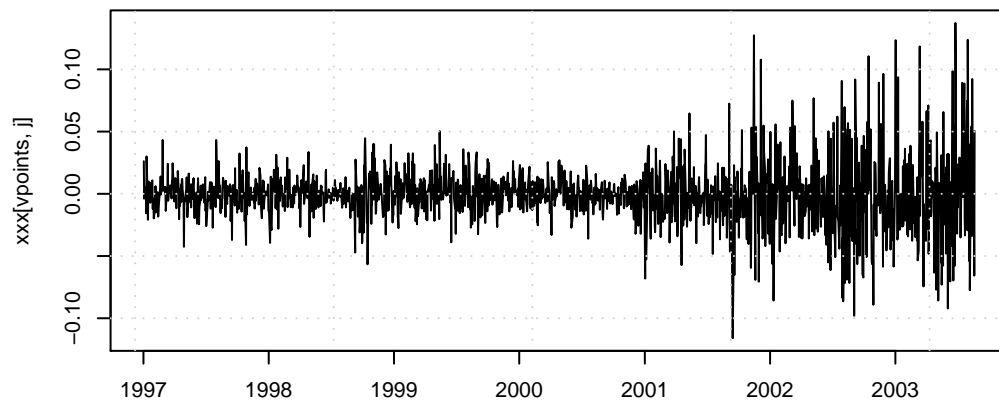


Figure 20: Time series for first PC of day-to-day changes in log forward rates.

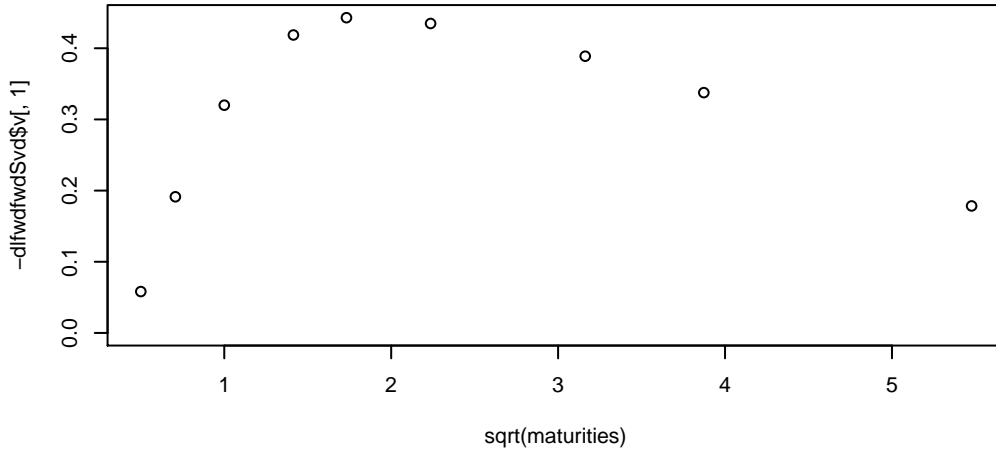


Figure 21: Loadings for first PC of scaled differences of log forward rates.

with a progressive decline in short-term rates. If the noise in the rates were roughly constant, we would expect the noise in their logarithms to increase as the rate decrease, as logarithmic changes are approximately the same as proportional changes. In this case, we might model

$$x_t = r_{t-1}(t_{i-1}, t_i) \log[r_t(t_{i-1}, t_i)/r_{t-1}(t_{i-1}, t_i)]$$

instead of the first usual first difference

$$\log[r_t(t_{i-1}, t_i)] - \log[r_{t-1}(t_{i-1}, t_i)] = \log[r_t(t_{i-1}, t_i)/r_{t-1}(t_{i-1}, t_i)].$$

We refer to this  $x_t$  as a *scaled* difference. The loadings for the first component for the scaled differences are shown in Figure 21. These loadings are qualitatively similar to those for the day-to-day changes in yields themselves (Figure 16). Figure 22 shows the corresponding time series component. Changes in volatility are much smaller than those in Figure 20, but are still visible.

The result of fitting a first-order autoregression to the time series is shown in Figure 23. The autoregressive coefficient `ar1` is small (0.07) but quite significant ( $P = 0.35\%$ ). The corresponding residuals are shown in Figure 24,

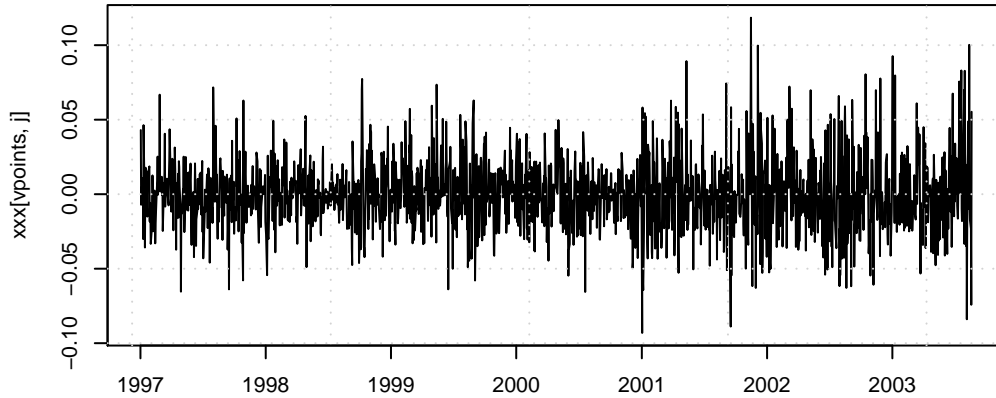


Figure 22: Time series for first PC of scaled differences log forward rates.

and are somewhat more zero-centered than the series in Figure 22. Figure 25 shows the GARCH(1, 1) fit to these residuals. The recursion parameter  $b_1$  is 0.93, again indicating a long time scale for volatility, and is highly significant, even though the variations in volatility are much smaller in Figure 20. The  $qq$ -plot of the residuals from this GARCH fit versus the Gaussian distribution is shown in Figure 26. The general curvature indicates skewness, and a slight ogive shape suggests long tails. The corresponding  $qq$ -plot against the  $t$ -distribution with 15 degrees of freedom, Figure 27, seems straighter overall. In summary, the first component time series for the scaled differences could plausibly be modeled as a first order autoregression with  $t_{15}$ -GARCH(1, 1) residuals.

```

Call:
arma(x = dlwdfwdSvd$u[, 1], order = c(1, 0, 0))

Model:
ARMA(1,0)

Residuals:
      Min       1Q   Median       3Q      Max
-1.146e-01 -1.242e-02  6.836e-05  1.458e-02  9.278e-02

Coefficient(s):
      Estimate Std. Error t value Pr(>|t|)
ar1      0.0701904  0.0240048   2.924  0.00346 **
intercept 0.0003400  0.0005766   0.590  0.55540
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Fit:
sigma^2 estimated as 0.000575,  Conditional Sum-of-Squares = 0.99,  AIC = -7989.65

```

Figure 23:  $AR(1)$  model for first component time series.

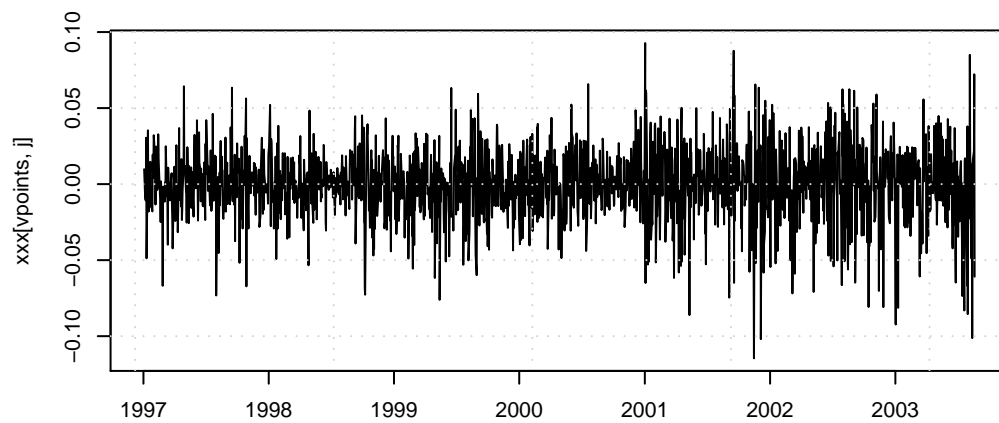


Figure 24: Residuals from ARMA fit.

```

Call:
garch(x = a)

Model:
GARCH(1,1)

Residuals:
      Min       1Q   Median       3Q      Max
-3.972461 -0.559544  0.003336  0.621991  4.362503

Coefficient(s):
      Estimate Std. Error  t value Pr(>|t|)
a0 9.812e-06   3.427e-06    2.863  0.00419 **
a1 5.952e-02   9.120e-03    6.527 6.72e-11 ***
b1 9.255e-01   1.214e-02   76.216 < 2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Diagnostic Tests:
Jarque Bera Test

data: Residuals
X-squared = 116.2346, df = 2, p-value < 2.2e-16

Box-Ljung test

data: Squared.Residuals
X-squared = 0.39, df = 1, p-value = 0.5323

```

Figure 25: GARCH model for ARMA residuals.

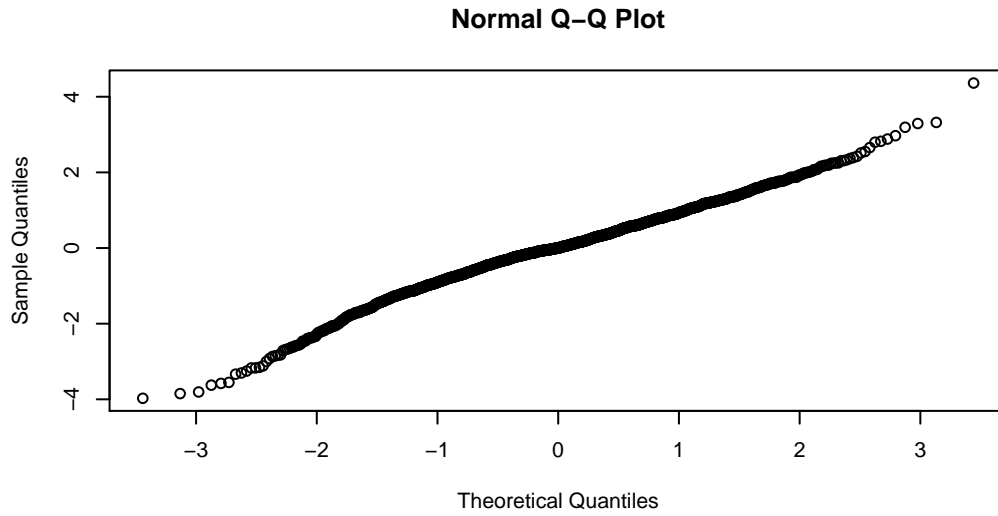


Figure 26: Gaussian  $qq$ -plot of GARCH residuals.

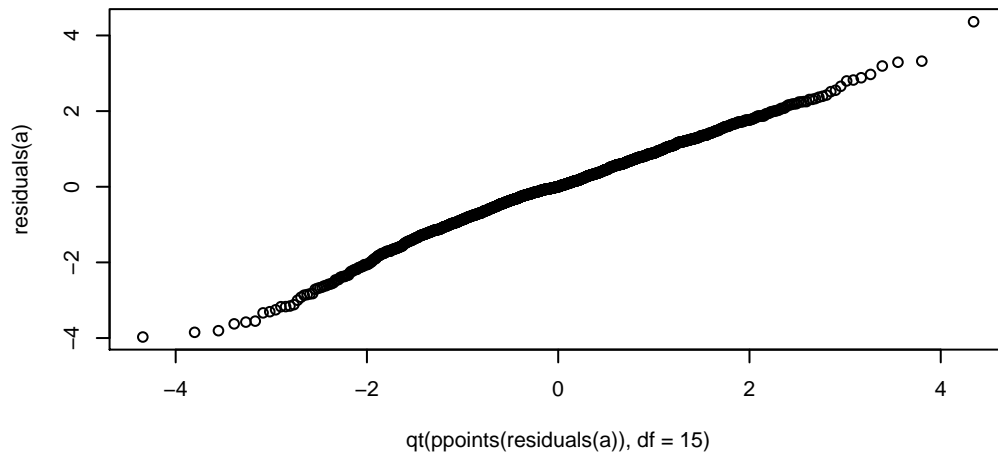


Figure 27:  $qq$ -plot of GARCH residuals against  $t$ -distribution with 15 degrees of freedom.