What is the most appropriate model for generating scenarios for daily foreign exchange rates?

Parker, John C.

relevant economics

June 2005

What Is The Most Appropriate Model For Generating Scenarios For Daily Foreign Exchange Rates?

John C. Parker

john.parker@relevanteconomics.com

June 2005

Abstract

This paper investigates the most appropriate model for generating scenarios for daily foreign exchange rates for a long history of a large number of daily exchange rates and finds: returns are not normal; a mean reversion model is rarely appropriate; sampling from historical returns (natural log differenced data) will capture the basic features of the mean of the return data but will ignore the autocorrelation in the mean and variance of returns; using a fat-tailed distributional assumption by matching the kurtosis of the historical data will capture the excess kurtosis of the data but similarly ignore these autocorrelations; a GARCH(1,1) model is in most cases sufficient to model time dependence of the conditional variance and will generate returns with excess kurtosis. In some cases an MA(1) - GARCH(1,1) model is required to capture residual autocorrelation, and in a few case more complicated ARMA(p,q) - GARCH(1,1) models are needed.

Keywords

ARIMA models; Exchange Rates; GARCH models; Risk Management; Scenarios; Time series; Vector Autoregression Models; Volatility forecasting.

Introduction

In the practice of risk management scenarios are a key input into understanding, measuring, and managing risk. Scenarios, like forecasting are views of the future. However scenario selection differs considerably from forecasting. A forecast is a prediction that a single scenario will occur. If you forecast the future and choose to only examine one scenario, the accuracy of a forecast becomes crucial. Yet no one is able to consistently forecast the future. So if the goal is to understand risk, a broad range of scenarios should be examined. The goal of selecting scenarios in risk management is to span the range of future events, not to forecast that any of these events will actually occur.

Historically summary statistics such as a covariance matrix have been used as inputs into the risk management process. However, scenarios have proved a much more robust and useful way of capturing the core input information required by a risk management framework. In risk methodologies where a summary statistic is used as the key information input, a set of scenarios is implied. By contrast, when using the scenario-based framework you must be explicit about your scenario choice.

The appropriate choice of scenarios, whether explicit or implicit, is the key factor that determines whether your risk analysis is adequate. Good risk management depends on the ability to generate relevant, forward-looking scenarios that properly represent the future.

Scenarios that embody correlated, consistent, simultaneous changes in market, credit and liquidity states naturally provide correlated, consistent risk output. Risk measures that link to these sources of risk can then be calculated.

What is the most appropriate model for generating scenarios for daily foreign exchange rates?

When generating scenarios using Monte Carlo methods for market risk analysis of foreign exchange rates using daily data some popular methods are:

- generating scenarios using the normality of returns assumption;
- a mean reversion model;
- sampling from historical returns (natural log differenced data) or historical daily differenced data;

- using a fat-tailed distributional assumption by matching the kurtosis of the historical data;
- using a GARCH model to model time dependence of the conditional variance and generate returns with excess kurtosis.

The first section documents the data used and the following sections roughly follow the outline of investigation the above listed issues. In addition, the question of whether a multivariate approach to exchange rates is warranted is investigated.

Data

This study looks at short-term exchange rate movements for market risk. While there are a few structural models to explain longer-term exchange rate movements, here we only use the history of an exchange rate as an explanation for future movements.

There is a long history of data (almost 30 years for some series) available from the Federal Reserve at http://www.federalreserve.gov/releases/H10/hist/ and http://www.federalreserve.gov/releases/H10/hist/thru89.htm. The exchange rates are based on data collected by the Federal Reserve Bank of New York from a sample of market participants for noon buying rates in New York for cable transfers payable in foreign currencies. There are data available on 23 exchange rates against the U.S. dollar. The data used in this study is as follows:

Country	Monetary unit	Mnemonic	Years of Data	Total Number of Observations	Start Date	End Date
Australia	Dollar	aud	29	7,585	04-Jan-71	06-Apr-01
Brazil	Real	brl	6	1,574	02-Jan-95	06-Apr-01
Canada	Dollar	cad	29	7,598	04-Jan-71	06-Apr-01
China, P.R.	Yuan	cny	19	5,031	02-Jan-81	06-Apr-01
Denmark	Krone	dkk	29	7,591	04-Jan-71	06-Apr-01
EMU member countries	Euro	eur	2	571	04-Jan-99	06-Apr-01
Hong Kong	Dollar	hkd	20	5,092	02-Jan-81	06-Apr-01
India	Rupee	inr	27	7,085	02-Jan-73	06-Apr-01
Japan	Yen	jpy	29	7,586	04-Jan-71	06-Apr-01
Malaysia	Ringgit	myr	29	7,570	04-Jan-71	06-Apr-01
Mexico	Peso	mxp	7	1,860	08-Nov-93	06-Apr-01
New Zealand	Dollar	nzd	29	7,576	04-Jan-71	06-Apr-01
Norway	Krone	nok	29	7,591	04-Jan-71	06-Apr-01
Singapore	Dollar	sgd	20	5,091	02-Jan-81	06-Apr-01
South Africa	Rand	zar	29	7,565	04-Jan-71	06-Apr-01
South Korea	Won	krw	19	4,976	13-Apr-81	06-Apr-01
Sri Lanka	Rupee	lkr	26	6,732	02-Jan-73	06-Apr-01
Sweden	Krona	swk	29	7,591	04-Jan-71	06-Apr-01
Switzerland	Franc	chf	29	7,592	04-Jan-71	06-Apr-01
Taiwan	Dollar	twd	16	4,108	03-Oct-83	06-Apr-01
Thailand	Baht	thb	19	5,011	03-Jan-81	06-Apr-01

United Kingdom	Pound	gbp	29	7,592	04-Jan-71	06-Apr-01
Venezuela	Bolivar	veb	6	1,573	02-Jan-95	06-Apr-01

Table 1: Data used in the study

Normalilty of Returns?

Scenarios for foreign exchange rates are often generated by assuming that returns are normally distributed. To be normally distributed the returns (defined as the difference in the natural log of the exchange rate) should have moments that match those of the normal distribution. We allow the second moment, the variance to vary, since we can scale the distribution by the standard deviation or square root of the variance. If we require matching of moments up to the fourth moment, then the distribution should have mean zero, skewness of zero, and kurtosis of 3. Defining y as the return based on the spot exchange rate S :

$$y_t = \ln(S_t) - \ln(S_{t-1})$$

As an example the returns for the Canadian Dollar are plotted below:

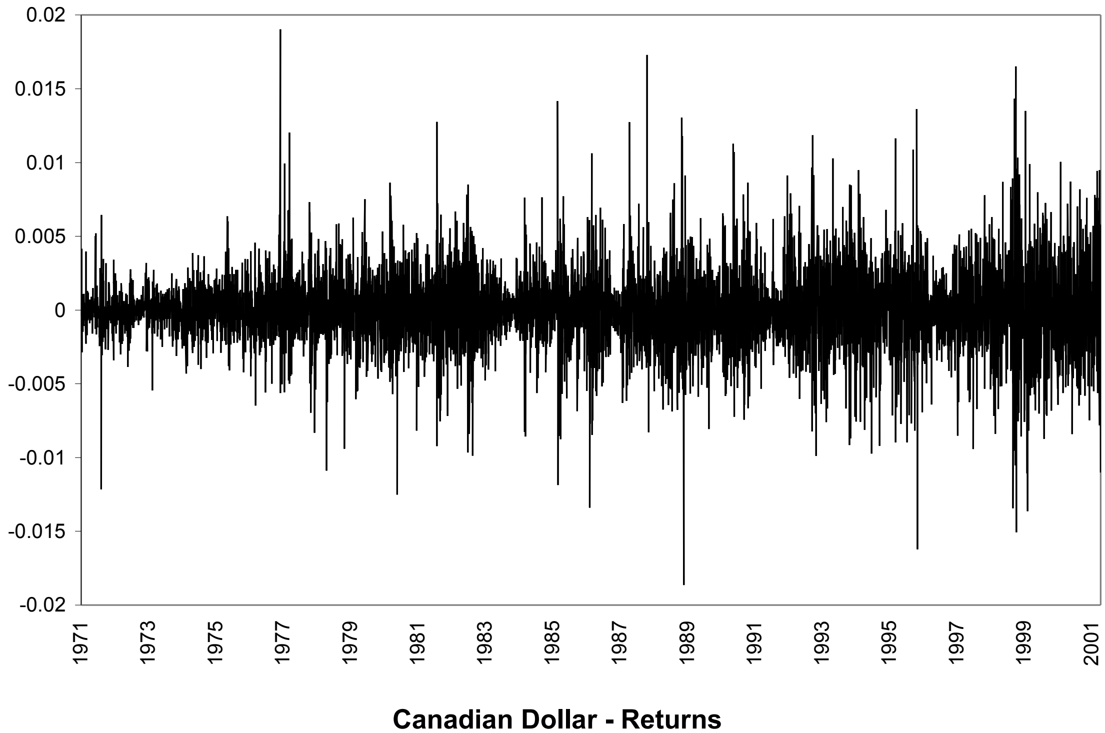

Figure 1: Natural log differences or returns in the Canadian vs. U.S. dollar exchange rate

Skewness is a measure of asymmetry of the distribution of the series around its mean. The skewness of a symmetric distribution, such as the normal distribution, is zero. Positive skewness means that the distribution has a long right tail and negative skewness implies that the distribution has a long left tail:

$$S = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^3$$

Kurtosis measures the peakedness or flatness of the distribution of the series. If the kurtosis exceeds 3, the distribution is peaked (leptokurtic) relative to the normal; if the kurtosis is less than 3, the distribution is flat (platykurtic) relative to the normal:

$$K = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - \bar{y}}{\hat{\sigma}}\right)^4$$

A test that combines the skewness and kurtosis is the Jarque-Bera test statistic for testing whether the series is normally distributed. The test statistic measures the difference of the skewness and kurtosis of the series with those from the normal distribution:

$$Jarque - Bera = \frac{N-k}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right)$$

where k is the number estimated coefficients. Under the null hypothesis of a normal distribution, the Jarque-Bera statistic is distributed as a χ^2 with 2 degrees of freedom. The probability tested is the probability that a Jarque-Bera statistic exceeds (in absolute value) the observed value under the null. A small probability value leads to the rejection of the null hypothesis of a normal distribution.

In all 23 cases, the probability is 0.00% and the assumption of normality is rejected.

Another form of normality test is to plot the quantiles of the series against the normal quantiles.

Here is a representative chart for the Brazilian Real showing the deviations of returns from normality in the tails of the distribution, shown as an S-shape versus the straight line that would be obtained for a normal distribution:

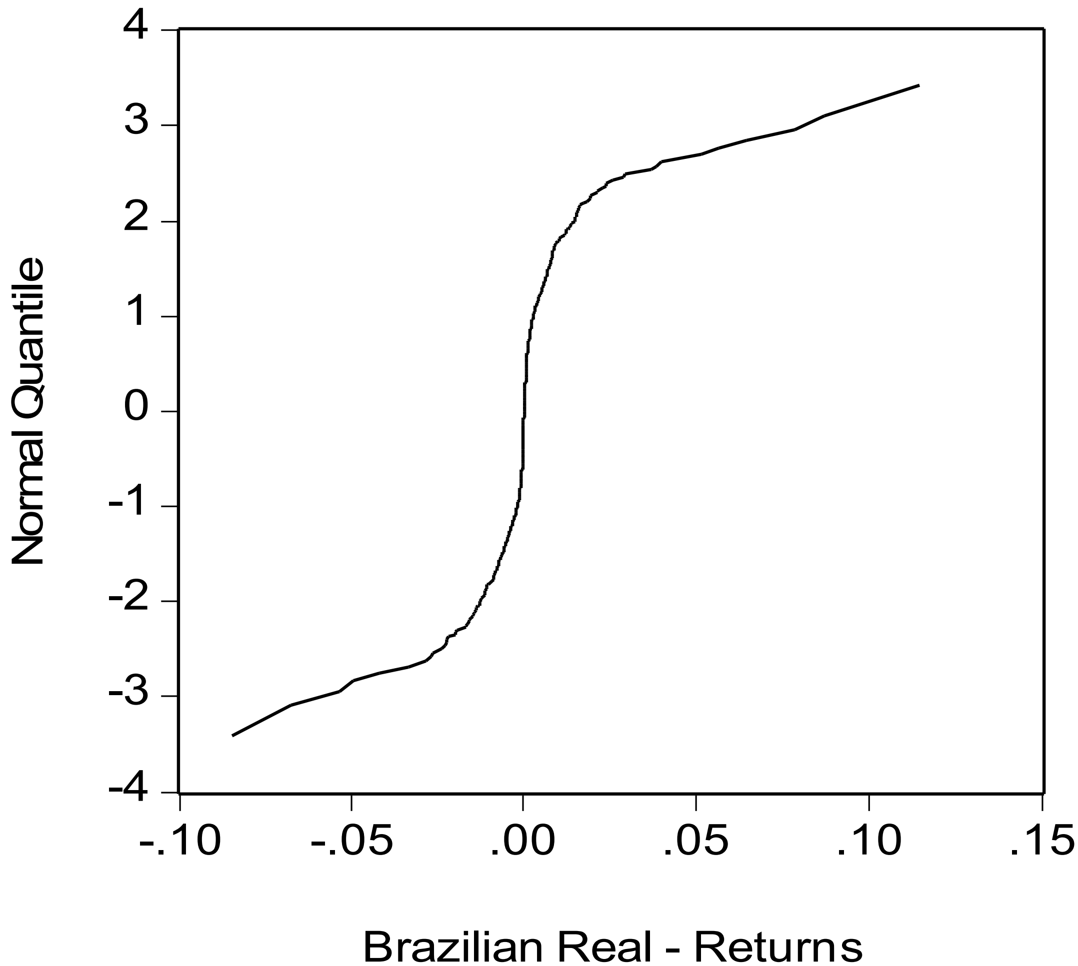

Figure 2: Quantile plot for returns of the Brazilian Real vs. U.S. dollar exchange rate

Mean Reversion or Unit Root?

Mean reversion is a popular assumption for foreign exchange rates in risk management. However there are some drawbacks in practice. Bernstein (1996) summarizes the problems:

> There are three reasons why regression to the mean can be such a frustrating guide to decision making. First, it sometimes proceeds at so slow a pace that a shock will disrupt the process. Second, the regression may be so strong that matters do not come to rest once they reach the mean. Rather, they fluctuate around the mean, with repeated, irregular deviations on either side. Finally, the mean itself may be unstable, so that yesterday's normality may be supplanted today by a new normality that we know nothing about. It is perilous to assume that prosperity is just around the corner simply because it has always been just around the corner.

Nonetheless, because it is a common assumption it should be investigated as a possible model for exchange rate scenario generation.

If we assume that exchange rates revert to the natural log of the level, the mean reversion process for a variable S, in continuous time, with mean λ, is as follows:

$$dS = \alpha(\lambda - \ln S)Sdt + \sigma Sdz$$

If we let:

$$x = \ln S$$

Then, we have:

$$dx = \alpha(\bar{x} - x)dt + \sigma dz$$

We can convert the above continuous time model into a discrete time model:

$$\Delta x = \alpha\bar{x}\Delta t - \alpha x\Delta t + \sigma\sqrt{\Delta t}\,\varepsilon$$

Now, we can regress x against Δx. The intercept will be $\alpha\bar{x}\Delta t$, and the slope will be $-\alpha\Delta t$.

The mean-reverting coefficients can be extracted from the estimated regression coefficients. Let ϖ_1 be the intercept, and φ_1 be the slope, then:

$$\alpha = -\frac{\varphi_1}{\Delta t}$$

$$\bar{x} = \frac{\varpi_1}{\alpha\Delta t}$$

The advantage of this model is that the application of Itô's Lemma results in parameters which are independent of S.

The volatility, σ, can be estimated by the standard error of the regression scaled by Δt.

Subsuming the Δt scaling for the frequency of the data, the parameter relationships, including the time dimension, the regression model can be written as:

$$\Delta \ln(S_t) = \varpi_1 + \varphi_1 \ln S_{t-1} + \eta_{1,t}$$

If we allow for the possibility of an autonomous trend:

$$\Delta \ln(S_t) = \varpi_1 + \varphi_1 \ln S_{t-1} + \beta_1 t + \eta_{1,t}$$

Note that one could also assume that the process is on the level of the exchange rate rather than the log. While in simulation this may cause the exchange rate to go negative, this possibility is also tested. So:

$$\Delta(S_t) = \varpi_2 + \varphi_2 S_{t-1} + \eta_{2,t}$$

or

$$\Delta(S_t) = \varpi_2 + \varphi_2 S_{t-1} + \beta_2 t + \eta_{2,t}$$

The alternative assumption is that the coefficient on the lagged exchange rate is -1, or the exchange rate has a unit root, or that:

$$\Delta \ln(S_t) = \varpi_1 + \eta_{1,t} \text{ or } \Delta(S_t) = \varpi_2 + \eta_{2,t}$$

The test of the mean reversion models versus the unit root model is a t-test on the parameter φ_1 or φ_2. The critical values are modified from the standard t-test with the values taken from Davidson and MacKinnon (1993, Table 20.1, page 708).

Using all of the data available the following table shows only the parameters in the mean reverting model are significant at the 10% level (*), 5% level (**), or 1% level (***):

Mnemonic	**Change in Level**	**Change in Log**	**Change in Level with Time Trend**	**Change in Log with Time Trend**
chf	***	*		
hkd	***	***	*	**

Table 2: Significance of Mean Reversion Model

These results suggest that for most currencies, the mean reverting model is rejected in favour of a unit root. Whether a currency has been fixed or pegged for a significant period of its history may affect this and we account for the exchange rate regime later in the paper.

The simple unit root test described above is correct if the series is an AR(1) process. If the series is correlated at higher order lags another test is required. The Augmented Dickey Fuller (ADF) and Phillips-Peron (PP) tests[1] adjust for higher-order serial correlation in the series. The ADF test simply adds more lags of the dependent variable to the right hand side of the regression equation. Another more general test of the unit root is the PP test that makes a correction to the t-statistic of the coefficient from the AR(1) regression to account for the serial correlation in the errors.

The PP test is done for levels, logs, differences, and returns (log differences). The following table shows whether the hypothesis of no unit root can be rejected at the 10% level (*), the 5% level (**) or the 1% level (***):

[1] All of the tests in this paper are easily computed in the econometric software package Eviews 4.0 (2001).

Mnemonic	Level	Natural Log of the Level	First Difference of the Level	Return (First Difference of the Natural Log)
chf	***	*	***	***
hkd	***	***	***	***
All others			***	***

Table 3: Phillips-Peron Test for a Unit Root

The PP test confirms that there is a unit root in the difference and log difference of all series, the level of chf, and for the level difference, returns, the level and log of the hkd.

Mean Reversion Over Some Sub-period?

Another possibility is that the mean reversion model is appropriate over some regime or sub-period of the data. Later we introduce variables to account for different regimes for exchange rates. A alternate, data-based approach, is to look at fixed size windows of one, two, five, and ten years of data and see how often the mean reverting model is selected over the unit root model, or the percentage of the time that the unit root model is rejected. In the table below the exceptions are shown along with the more representative results for the Italian Lira and the Malaysian Ringgit (full results are in xxx (2001)):

	1 year window		**2 year window**		**5 year window**		**10 year window**	
Mnemonic	**No. Tests**	**% Reject**	**No. Tests**	**% Reject**	**No. Tests**	**% Reject**	**No. Tests**	**% Reject**
eur	311	28.7%	51	79.4%	-	N/A	-	N/A
itl	***6,760***	***3.4%***	***6,500***	***2.3%***	***5,720***	***1.5%***	***4,420***	***0.7%***
mxp	1,600	17.4%	1,340	13.4%	560	18.4%	-	N/A
myr	***7,310***	***4.1%***	***7,050***	***2.9%***	***6,270***	***2.3%***	***4,970***	***1.6%***
veb	1,313	30.2%	1,053	27.3%	273	63.4%	-	N/A

Table 4: Test for Mean Reversion Over Sub-Period

Only a small number of times is the mean reversion model chosen over the unit root model such as: the Euro over one and two year windows; the Mexican Peso about 16% of the time; and the Venezuelan Bolivar over a longer data windows. In all these cases, there is less than 10 years of data and the results are therefore not conclusive.

Historical Sampling

In order for sampling of historical returns to be a valid approach to scenario generation, the level or returns and the variance of returns should be independent over time. Computing a correlogram, that is, the correlation of a series with its own lags, over some arbitrary set of lags, can test this assumption. For the level, and log level the one-day lag autocorrelation is always highly significant and averages 0.96 across the currencies. For the one-day difference and the log difference (return), in over half of the cases (53% and 62%) you can reject the hypothesis that the first autocorrelation is zero.

Given that differencing the natural log or the level of the series does not always allow us to reject the hypothesis that the autocorrelations are not statistically different from zero, it makes sense to try further differencing to remove any residual autocorrelations. In all cases for the log and level with the exception of the South Korean Won, further differencing increased the Q-statistic suggesting that differencing more than once was over-differencing and hence was introducing spurious autocorrelation. The conclusion then is that differencing once is appropriate but this does not always make the series independent over time for historical simulation generation. Testing all lags up to 36 allows the following conclusions to be made:

- Although probably not used, a simulation that assumes that the level is independent over time will be incorrect as there are no exchange rates that have no autocorrelation in the levels.
- Similarly, a simulation that assumes that the log of the exchange rate is independent over time will be incorrect as there are no log exchange rates that have no autocorrelation in the levels.
- A simulation that uses the difference of the level may be appropriate only for the following:
  - aud; cny; nzd; and veb (although modelling nzd in this way would ignore the apparent autocorrelation in the variance or squared returns).

- A simulation that uses returns (difference in the log) may be appropriate only for the following:
    - cny; lkr; and veb.

Fat-Tailed Distributions

Using a fat-tailed distributional assumption that matches the kurtosis of the historical data will capture the excess kurtosis of the data but will ignore the autocorrelation in the mean and variance of returns found above (62% of the return series and 78% of the squared return series had significant autocorrelation).

Variance Modeling Using GARCH

The analysis above on the squared returns suggests that there may be autocorrelation in the variance of exchange rate returns. A GARCH model captures this effect. In addition we introduce a variable to account for regime changes. The reason that this is done is that the regime dummy is required to ensure that the GARCH coefficients are correctly signed. Without the dummy some of the coefficients are implausible in that they may give negative variances in simulation. For the eur and twd a regime dummy was not included as the regime did not change over the sample period.

Using information from the International Monetary Fund on exchange rate regimes, we define the following variable for each exchange rate:

Regime	**Dummy Variable Value**
Pegged	1
Pegged exchange rates within horizontal bands	2
Crawling pegs	3
Limited Flexibility	4
Exchange rates within crawling bands	5
Managed floating with no pre-announced path for exchange rate	6
More Flexible: Adjusted according to a set of indicators	7
Managed floating	8
Independently floating	9

Table 5: Exchange rate regime dummy variables

In a GARCH model, the return equation is written as a function of a constant parameter (and perhaps with other exogenous terms such as our regime dummy variable) with an error term. σ_t^2 is the one-period ahead forecast of squared returns or the conditional variance. Also, the conditional variance equation is a function of three terms:

1. The average conditional volatility: ω.
2. News about volatility from the previous period, measured by the squared residual from the return equation: ε_{t-1}^2 (the ARCH term).
3. Last period's forecast of volatility: σ_{t-1}^2 (the GARCH term) (and perhaps with other exogenous terms such as our regime dummy variable).

In the GARCH(1,1) specification the return equation (1) and the conditional variance equation (2) are written:

$$y_t = \gamma + \varepsilon_t \quad \dots (1)$$

$$\sigma_t^2 = \omega + \beta\varepsilon_{t-1}^2 + \phi\sigma_{t-1}^2 \quad \dots (2)$$

The (1,1) in GARCH(1,1) refers to the presence of a first-order GARCH term, or a one period lag of volatility, and a first-order

ARCH term, or a one period lag of the squared residual from the return equation.

This model can be interpreted as predicting this period's volatility by a weighted average of a long-term average squared returns, the forecasted volatility from last period, and information about volatility observed in the previous period.

Modelling the variance of the returns of the exchange rates as a GARCH(1,1) process in all but two cases removes the autocorrelation in the variance series. For cad, gbp, krw, and sgd an ARCH Lagrange Multiplier T^*R^2 test (regressing the squared standardized residuals on lags of the same) rejects the hypothesis for no autocorrelation of the squared residuals (at the 1% level for cad, 0.1% level for gbp, 5% level for krw, and the 12% level for sgd). For these series, a GARCH(2,1) specification handles the squared residual autocorrelation.

In the GARCH(2,1) specification, compared with the GARCH(1,1) model, the conditional variance equation (2a) has an additional ARCH term:

$$y_t = \gamma + \varepsilon_t \quad \text{… (1a)}$$

$$\sigma_t^2 = \omega + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \varepsilon_{t-2}^2 + \phi \sigma_{t-1}^2 \quad \text{… (2a)}$$

There are tables in Xxx (2001) that show the significance of the GARCH model parameters. In general the parameters are highly significant for the variance equation and in some case the long-run returns (the constant from the mean equation) are significantly different from zero. For the model to be well-behaved we expect that the variance equation constant, ω **,** should positive (positive long-run variance), the ARCH(1) or β_1 , GARCH(1) or ϕ **,** should also be positive and the sum of ARCH(1), (ARCH(2) if included) and GARCH(1) should be less than 1, or $\phi + \sum_i \beta_i < 1$. While correctly signed in all case, for: cad; krw; nok; swk; twd; and zar the sum is greater than 1. This can give rise to potential explosive conditional variance estimates. This is checked again later, after an ARMA model for the returns is combined with a GARCH model for the variance.

The standardized residuals from these GARCH models are non-normal, as shown by the following tests (note that the errors have mean zero and standard deviation, on average of 0.7%, with a range of standard deviations from 0.3%-1.9%). For normality, the skewness and kurtosis should be 0 and 3. The Jarque-Bera statistic combines the excess skewness and kurtosis, over a normal distribution, and the probability shown is the probability that the residuals are normal):

Standardized Residuals from 24 Equations	**Skewness**	**Kurtosis**	**Jarque-Bera**	**Probability**
Mean	5.26	318.30	130,087,678	1.76E-12
Standard Deviation	12.84	667.74		
Minimum	-5.54	4.10	48	0.00E+00
Maximum	41.34	2694.66	2,033,845,226	4.04E-11

Table 6: Average normality of estimated GARCH residuals

Note that in the above table the average, minimum and maximum Jarque-Bera statistics are not related to the skewness and kurtosis shown, they are all calculated from a sample of 24 results. While the Jarque-Bera statistic combines the skewness and kurtosis the maximum Jarque-Bera statistic need not, for example, correspond to the maximum skewness if the resiuals had a very large kurtosis.

Combined Autoregressive Moving Average (ARMA) GARCH Analysis Models of Returns

Differencing the exchange rates once was found to be necessary to make the series stationary (that is to ensure the mean and variance are invariant over time). While there is a choice between changes in the levels and changes in the natural logs we use the latter since it has the interpretation of a daily return.

Modelling daily foreign exchange rate returns using ARMA models involves finding the optimal autoregressive lag length for the series and the optimal moving average lag length for the error term. The terminology for the models is ARMA(p,q), where p is the autoregressive (AR) length and q id the moving average (MA) length.

Model selection for an ARMA model is usually based on an information criterion. An information criterion provides a measure of the goodness of fit with penalties for the number of parameters in the model. This results in a specification of the model that is parsimonious. The various information criteria differ in how to strike the balance between fit and parsimony: Akaike information criterion (AIC); Schwarz criterion (SIC); Hannan-Quinn criterion (HQIC). The information criteria are all based on minus 2 times the average log likelihood function, adjusted by a penalty:

$$AIC = -\frac{2l}{n} + \frac{2k}{n}$$

$$SIC = -\frac{2l}{n} + \frac{k \log n}{n}$$

$$HQIC = -\frac{2l}{n} + \frac{2k \log(n)}{n}$$

where k is the number of estimated parameters, n is the number of observations, and l is the value of the log likelihood function.

In this section we put together the ARMA and GARCH analysis for returns. We tried first to re-estimate the GARCH with the best ARMA model, based in the SIC, for the mean equation. Then we jointly estimated the ARMA and GARCH orders by iterating over all combinations and choosing the model with the best SIC. To test the specification of the ARMA-GARCH model, all combinations of the ARMA and GARCH orders are estimated and compared in terms of the log likelihood fit and their parsimony. The best model specifications chosen, based on the SIC, from all ARMA-GARCH models from ARMA(0,0), to ARMA(3,3)-GARCH(2,2). In some cases high order ARMA terms were tried as the maximum order model was chosen with the preliminary screen.

This approach was used as a screening tool that narrowed down the models under consideration. This is because the significance of the parameters does not enter into the SIC and therefore the final specification was chosen by a combination of maximization of the log likelihood function, subject to parsimony and significance constraints.

Using returns, or the change in the natural log, as the basis for the models, the following are the models selected:

Mnemonic	**ARMA-GARCH with Regime Dummy**
aud	0,3,1,1
brl	0,0,1,1
cad	0,0,1,1
chf	0,0,1,1
cny	0,0,1,1
dkk	0,0,1,1
eur	0,0,1,1
gbp	0,0,1,1
hkd	1,2,1,1
inr	0,0,1,1
jpy	0,1,2,1
krw	0,0,1,1
lkr	0,1,2,1
mxp	0,0,2,1
myr	4,2,2,1
nok	0,0,2,1
nzd	0,3,1,1
sgd	0,0,2,1
swk	0,0,1,1
thb	2,2,2,0
twd	2,1,1,1
veb	0,0,1,1
zar	0,3,2,1

Table 7: Estimated ARMA-GARCH model orders

In 48% of the models a simple GARCH(1,1) model of returns is chosen. In 22% of the models a GARCH(2,1) model is chosen. In the remainder of cases a more complicated ARMA-GARCH model is chose, but with the GARCH form relatively simple.

In all cases the GARCH parameter restrictions are upheld. As before, as in the case of simple GARCH, for ARMA-GARCH, the variance equation constant, ω **,** should be positive (positive long-run variance), the ARCH(1) or β_1 , GARCH(1) or ϕ **,** should also be positive and the sum of ARCH(1), (ARCH(2) if included) and GARCH(1) should be less than 1, or $\phi+\sum_i \beta_i<1$. Now, in all cases, the parameter restrictions are valid allowing volatility estimates that are positive and non-explosive in simulation.

The standardized residuals from these GARCH models are non-normal, as shown by the following tests (note that the errors have mean zero and standard deviation, on average of 0.7%. With a range of standard deviations from 0.3%-1.9%). For normality, the skewness and kurtosis should be 0 and 3. The Jarque-Bera statistic combines the excess skewness and kurtosis, over a normal distribution, and the probability shown is the probability that the residuals are normal):

Standardized Residuals from 23 Equations	**Skewness**	**Kurtosis**	**Jarque-Bera**	**Probability**
Mean	5.44	328.63	141,291,318	1.76E-12
Standard Deviation	13.00	693.17		
Minimum	-5.55	4.10	48	0.00E+00
Maximum	42.68	2851.72	2,280,000,000	4.04E-11

Table 8: Average normality of estimated ARMA-GARCH residuals

Note that in the above table the average, minimum and maximum Jarque-Bera statistics are not related to the skewness and kurtosis shown, they are all calculated from a sample of 24 results. While the Jarque-Bera statistic combines the skewness and kurtosis the maximum Jarque-Bera statistic need not, for example, correspond to the maximum skewness if the resiuals had a very large kurtosis.

Even the Euro equation, which has the lowest combined skewness and kurtosis of the equations (0.45 and 4.10), produces a Jarque-Bera statistic that rejects normality at the 4.04×10^{-9} % level.

The excess kurtosis of suggesting that a model such as GARCH(1,1) with t distributed errors might be used for scenario generation. Estimation with of a GARCH model with t-distributed errors is an area for further research.

Multivariate Analysis?

It has been established that there is an autoregressive structure in the variance of the series. This section tests whether there is a need for simultaneous modelling of this structure, that is, whether there is any correlation in the errors from the filtering done thus far. While there is no formal test for whether multivariate analysis is appropriate there are a couple of possible screens that can be applied.

First, the correlation of the residuals from the ARMA-GARCH models can be calculated and examined for related errors.

Second a vector autoregression (VAR) model can test for the possibility that lagged log changes in one exchange rate can affect another rate. Shown below are the lags, from one to five, of a VAR model, that are significant at the 5% level (note that gbp(-3) is the third period (business day) lag of the gbp return):

aud	**brl**	**cad**	**chf**	**cny**	**eur**	**gbp**	**hkd**	**inr**	**jpy**	**krw**
mxp(-3)	chf(-3)	aud(-1)	cny(-5)	gbp(-3)	cny(-5)	inr(-3)	aud(-4)	chf(-5)	aud(-4)	veb(-4)
myr(-4)	eur(-3)	brl(-2)	hkd(-3)	hkd(-3)	hkd(-3)	krw(-1)	hkd(-1)	eur(-5)	brl(-1)	zar(-3)
nzd(-1)	thb(-3)	cad(-3)	jpy(-5)	inr(-4)	inr(-3)	krw(-4)	nzd(-4)	inr(-1)	krw(-4)	
sgd(-2)		cad(-4)	krw(-2)	swk(-1)	jpy(-4)		thb(-3)	inr(-4)	zar(-3)	
sgd(-3)		eur(-5)	nzd(-1)	veb(-5)	krw(-2)		zar(-4)	krw(-1)		
veb(-1)		gbp(-3)	nzd(-4)		krw(-4)			krw(-3)		
		hkd(-2)	sgd(-5)		nzd(-4)			mxp(-1)		
		krw(-3)						nok(-1)		
		lkr(-1)						sgd(-4)		
		twd(-4)						sgd(-5)		
								thb(-3)		

lkr	**mxp**	**myr**	**nok**	**nzd**	**swk**	**sgd**	**thb**	**twd**	**veb**	**zar**
chf(-4)	aud(-2)	myr(-1)	aud(-5)	aud(-1)	hkd(-2)	hkd(-2)	inr(-3)	aud(-3)	gbp(-1)	brl(-5)
cny(-2)	chf(-5)	myr(-2)	brl(-1)	chf(-3)	krw(-2)	inr(-1)	krw(-5)	cad(-2)	krw(-5)	cad(-1)
eur(-2)	eur(-3)	myr(-3)	cny(-5)	nok(-3)	krw(-4)	krw(-2)	sgd(-2)	chf(-3)	twd(-1)	cny(-3)
eur(-4)	eur(-5)	myr(-4)	hkd(-2)	nzd(-1)	myr(-2)	thb(-2)	thb(-2)	krw(-1)	veb(-1)	hkd(-4)
inr(-1)	gbp(-3)	myr(-5)	inr(-3)	nzd(-4)	sgd(-1)	veb(-5)		twd(-1)		krw(-4)
jpy(-2)	jpy(-2)		jpy(-4)	sgd(-2)				twd(-2)		myr(-4)
lkr(-2)	jpy(-3)		krw(-2)	thb(-1)						nok(-1)
mxp(-1)	krw(-2)		krw(-4)	thb(-2)						nzd(-1)
myr(-2)	mxp(-1)		mxp(-3)	thb(-4)						
myr(-3)	nok(-5)			thb(-5)						
myr(-4)	swk(-5)									
myr(	sgd(-									

-5)	2)									
swk(- 2)	sgd(- 3)									
	sgd(- 5)									
	thb(- 2)									
	twd(- 2)									
	veb(- 3)									
	zar(- 2)									
	zar(- 4)									

Table 9: VAR models for exchange rate returns

Similarly, one can run the same analysis using squared returns rather than returns to test for the need for multivariate GARCH:

aud	**brl**	**cad**	**chf**	**cny**	**eur**	**gbp**	**hkd**	**inr**	**jpy**	**krw**
sgd(-2)	brl(-1)	brl(-1)	gbp(-3)	gbp(-3)	gbp(-3)	cad(-5)	gbp(-3)	cad(-5)	brl(-2)	aud(-4)
	brl(-2)	brl(-2)	gbp(-4)	inr(-4)	gbp(-4)	chf(-5)	hkd(-5)	chf(-5)	chf(-3)	chf(-2)
	cad(-2)	brl(-3)	krw(-2)	lkr(-2)	hkd(-3)	eur(-5)	jpy(-3)	cny(-3)	chf(-5)	cny(-3)
	jpy(-1)	brl(-4)	krw(-3)	mxp(-1)	hkd(-4)	hkd(-3)		eur(-5)	eur(-5)	cny(-4)
	krw(-3)	gbp(-1)	krw(-4)	mxp(-3)	krw(-2)	krw(-3)		gbp(-5)	gbp(-2)	eur(-2)
	mxp(-2)	gbp(-3)	zar(-4)	mxp(-5)	krw(-4)	krw(-4)		krw(-2)	jpy(-1)	hkd(-5)
	sgd(-4)	inr(-5)		zar(-5)	zar(-5)	krw(-5)		zar(-2)	jpy(-4)	krw(-1)
	veb(-1)	jpy(-2)						zar(-4)	krw(-5)	krw(-4)
	zar(-4)	krw(-3)							mxp(-2)	lkr(-1)
		nok(-3)							sgd(-5)	lkr(-3)
		nzd(-3)							veb(-1)	swk(-1)
		twd(-2)								swk(-5)
		zar(-4)								

lkr	**mxp**	**myr**	**nok**	**nzd**	**swk**	**sgd**	**thb**	**twd**	**veb**	**zar**
chf(-4)	hkd(-4)	gbp(-2)	aud(-5)	cny(-1)	aud(-4)	sgd(-2)	jpy(-1)	aud(-2)	brl(-1)	brl(-5)
cny(-2)	inr(-3)	mxp(-3)	eur(-3)	nzd(-1)	krw(-4)		nzd(-3)	aud(-4)	brl(-3)	cad(-3)
eur(-4)	jpy(-1)	mxp(-4)	hkd(-4)	nzd(-2)	mxp(-3)		thb(-1)	gbp(-2)	krw(-5)	cad(-5)
gbp(-5)	jpy(-2)	myr(-1)	krw(-2)	nzd(-3)	zar(-5)		thb(-5)	hkd(-3)	thb(-4)	chf(-4)
lkr(-2)	krw(-2)	myr(-2)	krw(-4)	nzd(-4)			veb(-4)	hkd(-5)	veb(-1)	chf(-5)
myr(-3)	krw(-3)	myr(-3)		sgd(-2)				krw(-4)		eur(-5)
myr(-4)	mxp(-2)	myr(-4)		thb(-1)				thb(-3)		gbp(-3)
myr(-5)	nok(-4)	myr(-5)		thb(-2)				twd(-1)		gbp(-4)
	swk(-2)	swk(-2)		thb(-4)				twd(-3)		lkr(-1)
	swk(-4)	zar(-3)						twd(-4)		mxp(-2)

	zar(- 3)									mxp(-4)
	zar(- 4)									zar(- 1)

Table 10: VAR models for exchange rate squared returns

These tables suggest groups of exchange rates that should probably be considered together. Thus someone with exposure to the New Zealand and Australian exchange rates (vis-à-vis the US dollar) should probably model the exchange rates together as there appears to be a two-way effect between these currencies. On the other hand, someone with exposure to the Pound and the Yen can probably model them independently.

To put these two tables in perspective the average number of significant lags is 6-7 which is 6.2-6.8% of the 110 combinations of lags tested.

As a further test for multivariate analysis, one can also calculate the correlation of the residuals from the final ARMA-GARCH models to see whether they are correlated.[2]

Forecasts

Forecasts were generated for a couple of variables to test the models. The estimation was done to April 6, 2001. An extra 39 observations, up to June 1, 2001 were available to test the

[2] For this data set this is difficult, since the errors from the GARCH model are not "dated". That is in order to create the continuous sample required for maximum likelihood GARCH estimation they were dumped to a vector to take out N/A's and then re-read into "dated" vectors that do not correspond to their actual dates. Thus, the errors in a variance-covariance matrix do not line-up.

models. Out of sample forecasts of gbp, jpy, and inr were made, and their confidence intervals (plus and minus one and two standard deviations – sd) plotted against the actuals and two years of history:

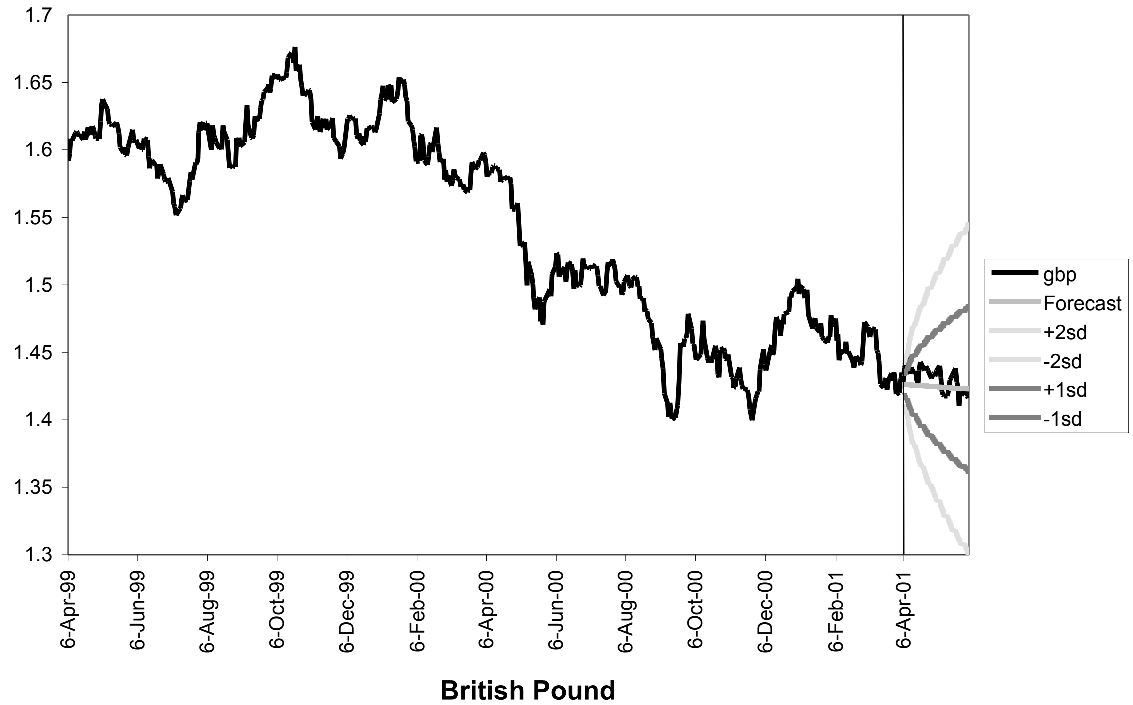

Figure 3: Forecast and standard error bands for scenarios generated for the British Pound vs. U.S. dollar exchange rate

For the gdp exchange rate a hundred scenarios were generated using the models and the 1%, 5%, 10%, 90%, 95%, and 99% percentile scenarios are also plotted:

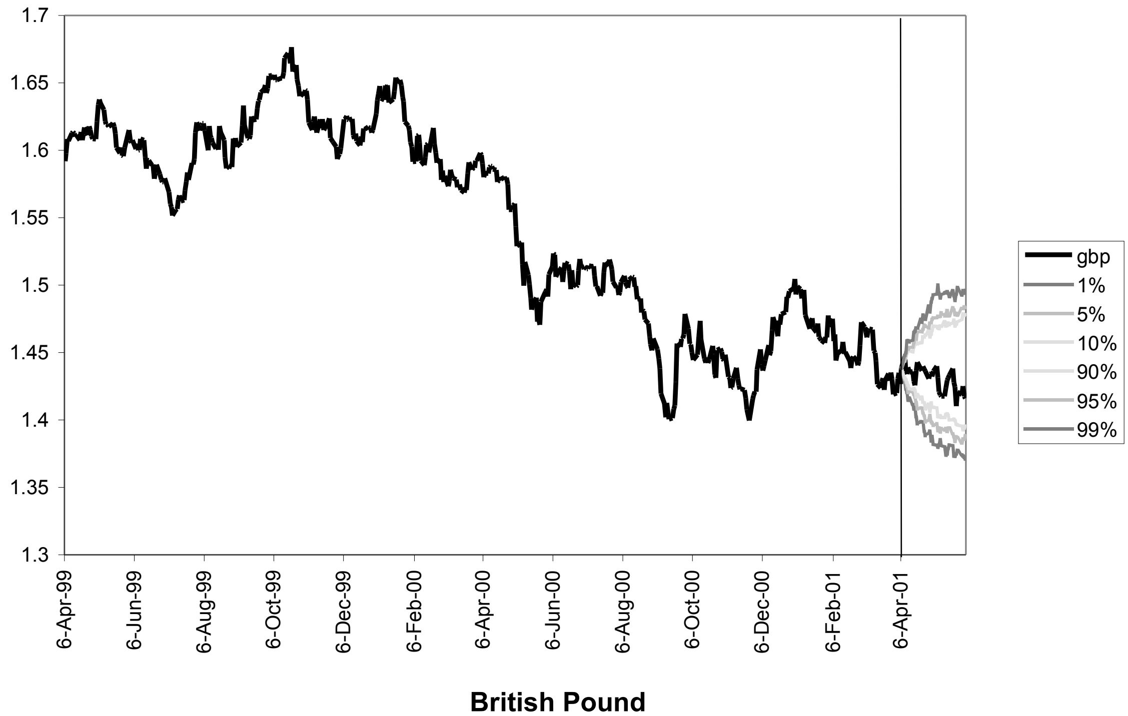

Figure 4: Forecast and four scenarios generated for the British Pound vs. U.S. dollar exchange rate

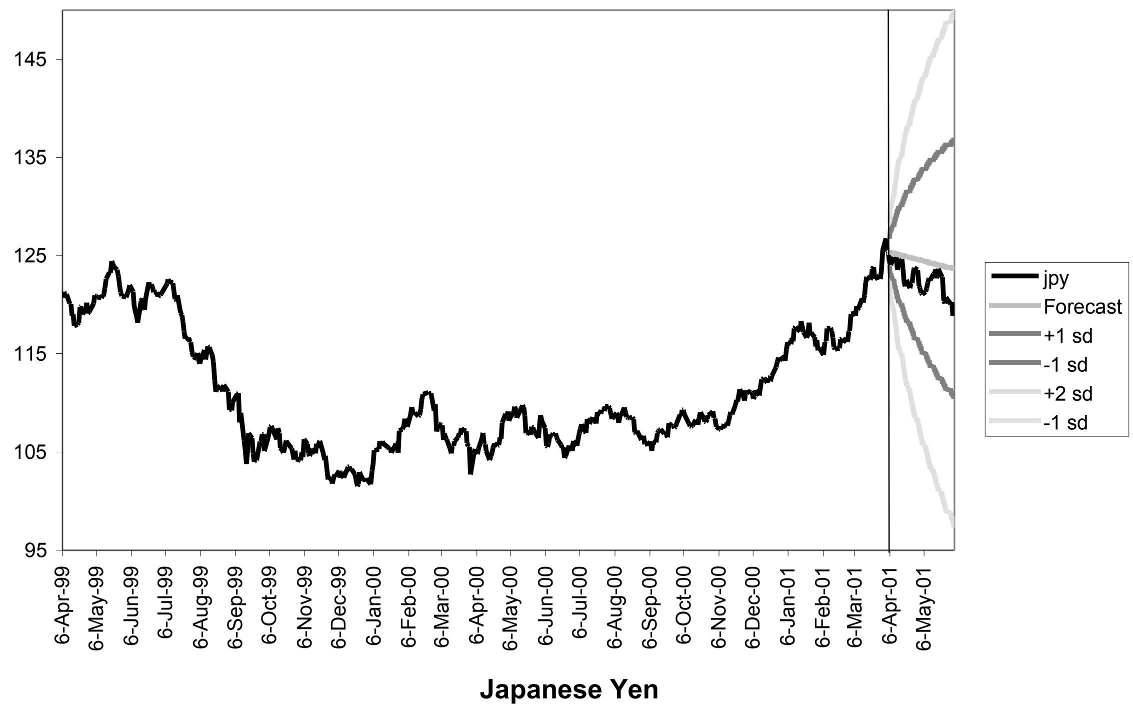

Figure 5: Forecast and standard error bands for scenarios generated for the Japanese Yen vs. U.S. dollar exchange rate

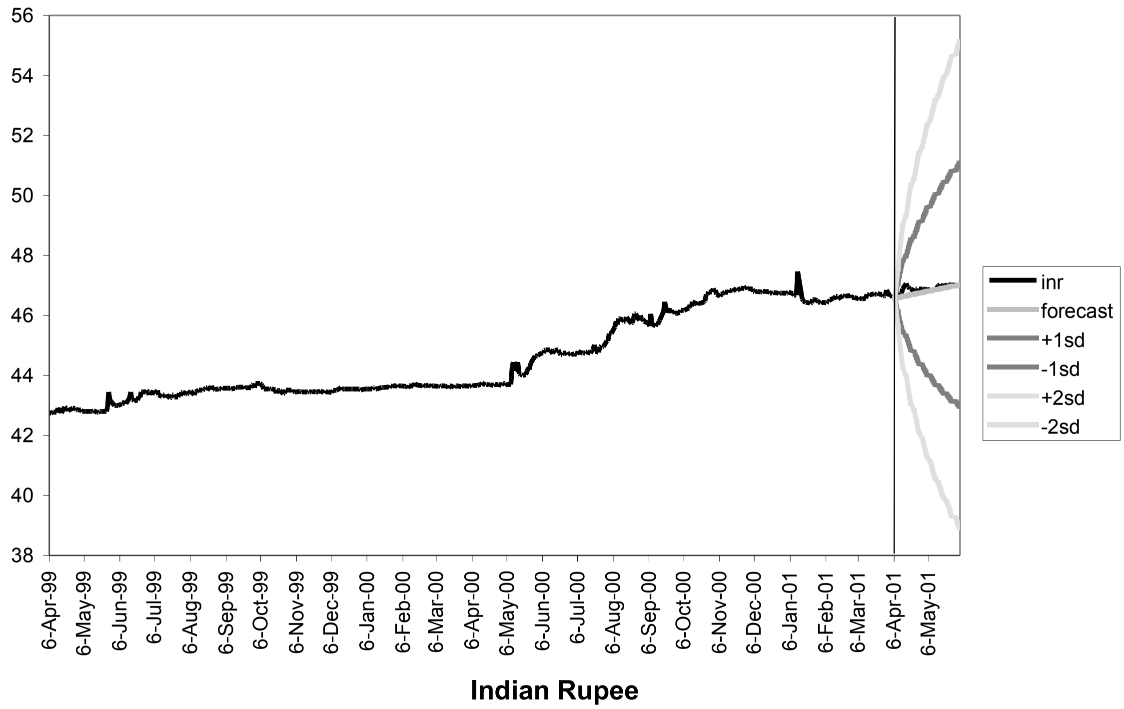

Figure 6: Forecast and standard error bands for scenarios generated for the Indian Rupee vs. U.S. dollar exchange rate

In all cases the models capture the trend and distribution in exchange rates. The scenarios generated from these models are quite plausible. While the error bands and extreme scenarios encompass the actual data, the bands are quite large. It should be noted though that these are mutli-step out of sample forecasts without any updating of parameters or use of actuals in the scenarios.

Conclusions

Normality of Returns

In all cases, the assumption of normality is resoundingly rejected.

Historical sampling

- A simulation that uses the difference of the level may be appropriate only for the following:
  - aud; cny; nzd; and veb (although modelling nzd in this way would ignore the apparent autocorrelation in the variance or squared returns).
- A simulation that uses returns (difference in the log) may be appropriate only for the following:
  - cny; lkr; and veb.

Mean Revertion

As Peter Bernstein (Op. Cit.) suggests: "The trick is to flexible enough to recognize that regression to the mean is only a tool; it is not a religion with immutable dogma and ceremonies. Used to make mechanical extrapolations of the past … regression to the mean is little more than mumbo-jumbo. Never depend on it to come into play without constantly questioning the relevance of the assumptions that support the procedure" (pp. 185-186). Mean reversion is strongest for fixed exchange rate regimes but moves can be extreme when adjustments come. While mean reversion may exist over some periods it is not a stable or reliable model.

Fat-Tailed Distributions

Using a fat-tailed distributional assumption that matches the kurtosis of the historical data will capture the excess kurtosis of the data but will ignore the autocorrelation in the mean and variance of returns that is found in the data.

GARCH Models

Modelling the variance of the returns of the exchange rates as a simple GARCH(1,1) process is most cases is enough to remove the autocorrelation in the variance series. cad, gbp, lkr, and sgd are better represented by a GARCH(2,1) model. Standardized residuals from the GARCH models are non-normal. In some cases a few case more complicated ARMA - GARCH models are needed. ARMA-GARCH models generate reasonable and plausible scenarios.

Multivariate Analysis

Tables are provided in the text for those series that should be combined into blocks and modeled simultaneously.

References

Bernstein, P.L. (1996). *Against the Gods - the Remarkable Story of Risk*, Wiley.

Davidson, R. & MacKinnon, J.G. (1993), *Estimation and Inference in Econometrics*, Oxford University Press.

Eviews 4.0 (2001), Quantitative Micro Software, LLC, Irvine CA., www.eviews.com

Parker, John C, (2001), What Is The Most Appropriate Model For Generating Scenarios For Daily Foreign Exchange Rates?, *Algorithmics Inc. Research Paper Series*: 01-01 July 21, 2001 – available from the author john.parker@relevateconomics.com .