A Statistical Arbitrage FX Trading System Based on Short Term FX Volatility Swings Forecasting with Institutional Data on JPY based Investment flows into US Markets

(Internal flow data of daily frequency maintained by a New York based large custodian bank for investment flows into US equities and fixed income instruments by Japanese traders/funds employed for this study)

Pavan Gadiraju

First Exposition, June 2009

Abstract:

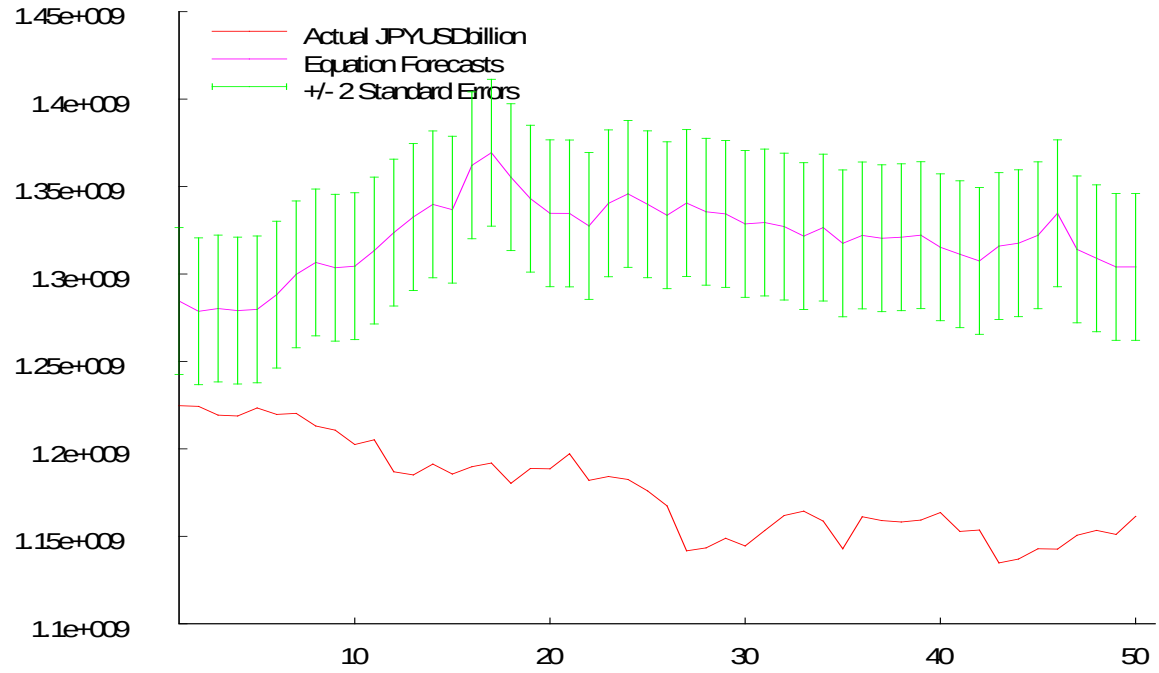
In this paper I explore the informational content in cross-currency flow maintained by large custodian banks with an objective to design a statistical arbitrage trading system that could exploit such information. After an initial simple test involving one-step ahead forecasts for JPYUSD FX pair with lagged I-Flow data series and concluding that such forecasts don't measure up to a simple AR(1) model's forecasts involving the FX pair time series itself, I introduce a 15-day moving standard deviation variables based off the I-Flow time series with a 5-day lag to the one-period I-Flow forecasting model to discover a considerable improvement in forecasts to the original model though still not bettering the simple AR(1) regression model of the FX time series. With the information from the initial tests at hand, we move on to explore the possibility of designing a system to forecast the swings observed in the 15-day moving standard deviation series of the JPYUSD FX pair. A partial dynamic equilibrium regression system involving a transformation of the individual I-Flow series similar to the FX pair series (15-Day moving standard deviation) is then experimented with to capture long-run stable 5-period ahead forecast for the 15-day moving standard deviation swings in the JPYUSD FX pair. Finally, I propose two models to exploit the volatility swing forecasting system; The first model involving volatility trades for the FX pair that could exploit an expected upswing in the short term volatility of the FX pair, and the second model involving an overlay of the swing forecasting system over traditional trend-forecasts involving technical rules to capture profitable long/short trades of the FX pair. **Data between May 2007 and May 2009 was employed for the exercise.**

Initial Tests for one-period forecasts: Period 450 – 550

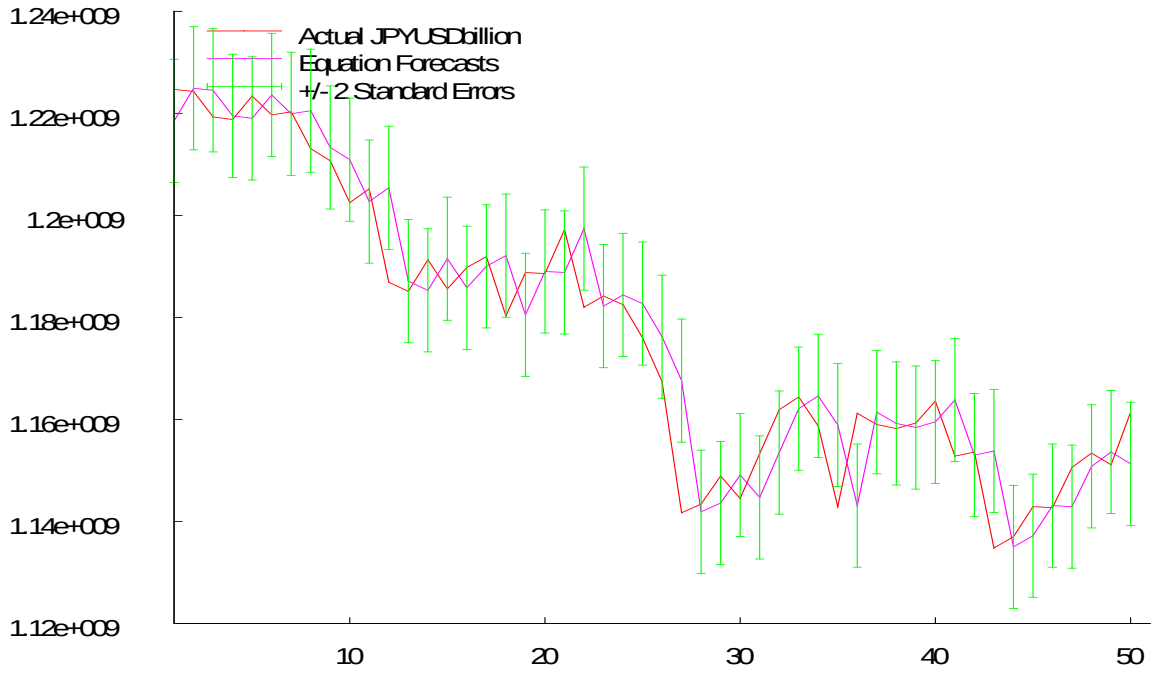
Stable model designed after progressively dropping insignificant variables based on P-Values.

- **********
- TSM4.29.22-01-09 Run 141 at 12:03:31 on 7-06-2009
- Data file is C:\Users\quantabaye\Documents\JPYUSDFlowV1.csv
- ---
- Dependent Variable is JPYUSDbillion
- 101 observations (450-550) used for estimation
- with 449 pre-sample observations.
- Estimation Method: Ordinary Least Squares
    - Estimate Std. Err. t Ratio p-Value
- JPFI:10+YearsCumulative(-1) -0.0007 0.00015 -4.678 0
- JPFI:5-10YearsCumulative(-1) -0.00056 0.00011 -5.061 0
- JPFI:2-5YearsCumulative(-1) -0.00077 9e-005 -8.598 0
- JPFI:1-2YearsCumulative(-1) -0.00136 0.0006 -2.267 0.026
- JPFI:6-12MonthsCumulative(-1) -0.00309 0.00072 -4.296 0
- JP EQ/FI Total Cumulative(-1) 0.00065 4e-005 16.338 0
    - Log Likelihood = -1843.04
    - Schwarz Criterion = -1856.89
- Hannan-Quinn Criterion = -1852.22
    - Akaike Criterion = -1849.04
    - Sum of Squares = 4.18596e+016
        - R-Squared = 0.4269
    - R-Bar-Squared = 0.3967
        - Residual SD = 2.09892e+007
    - Residual Skewness = -0.4514
    - Residual Kurtosis = 3.342
    - Jarque-Bera Test = 3.9226 {0.141}
- Box-Pierce (residuals): Q(12) = 141.649 {0}
- Box-Pierce (squared residuals): Q(12) = 83.3808 {0}
- Cointegration Test (ADF,0 lags) = .NaN {?}
- Cointegration Test (PP) = -2.3247 {?}*
- Lag length for ADF test selected by Akaike criterion.
- Covariance matrix from robust formula.
- PP bandwidth = 3 (Parzen kernel.)
- ...Run completed in 0.33

1-Period forecasts based on above regression model for 50-Periods:



Similar 1-period forecasts of an AR(1) model of the JPYUSD time series for 50 periods: JPYUSD regressed against its one-period lag



Evidence from above that the I-Flow regression system fails in predicting FX pair movements while the AR(1) model, as expected, tracks the FX pair movements closely though offering no advantage for generating profitable trades.

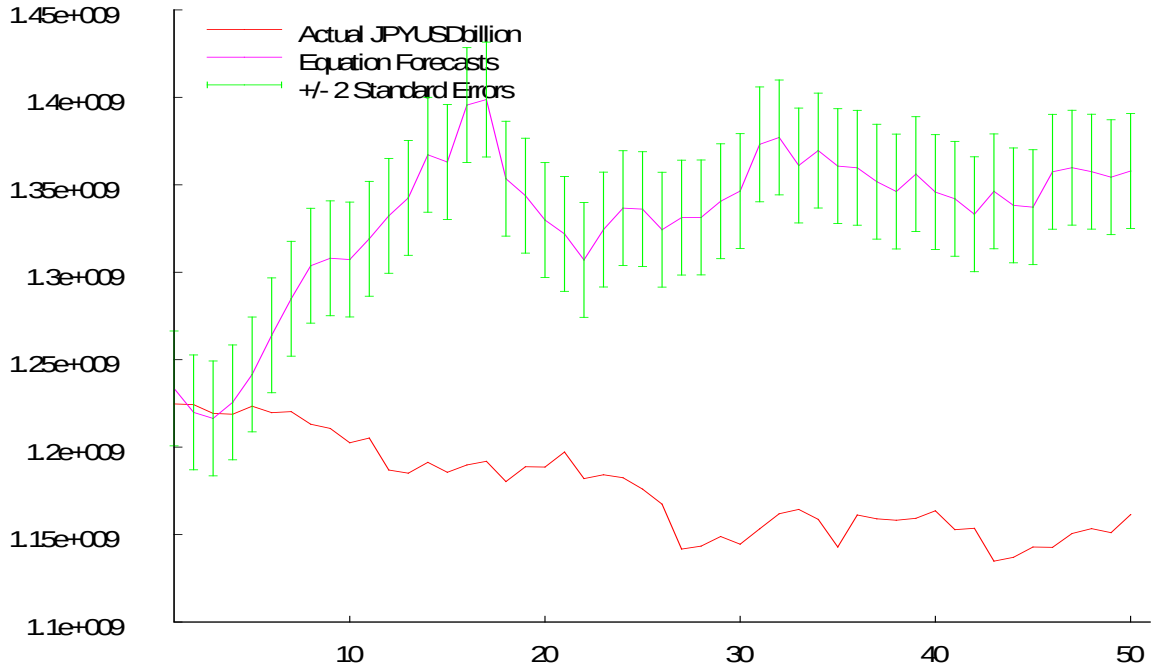
Second Model: 15-Day moving standard deviation series with a 5-period lag introduced to the above I-Flow regression system and a stable model designed after a similar exercise involving dropping insignificant variables based on P-Values.

- **********
- TSM4.29.22-01-09 Run 149 at 12:12:54 on 7-06-2009
- Data file is C:\Users\quantabaye\Documents\JPYUSDFlowV1.csv
- ---
- Dependent Variable is JPYUSDbillion
- 101 observations (450-550) used for estimation
- with 449 pre-sample observations.
- Estimation Method: Ordinary Least Squares
- Estimate Std. Err. t Ratio p-Value
- JPFI:10+YearsCumulative(-1) -0.00084 0.00022 -3.812 0
- JPFI:5-10YearsCumulative(-1) -0.00078 0.0001 -7.776 0
- JPFI:2-5YearsCumulative(-1) -0.00107 9e-005 -11.894 0
- JPFI:6-12MonthsCumulative(-1) -0.00557 0.00094 -5.931 0
- JPFI:0-6MonthsCumulative(-1) 0.0002 9e-005 2.229 0.028
- JP EQ/FI Total Cumulative(-1) 0.00068 5e-005 13.527 0
- JPFI:5-10YearsCumulative15DaySdev(-5) 0.0019 0.00043 4.419 0
- JPFI:2-5YearsCumulative15DaySdev(-5) 0.00077 0.00037 2.077 0.041
- JPFI:6-12MonthsCumulative15DaySdev(-5) 0.00441 0.00126 3.499 0.001
- JPEQ/FITotalCumulative15DaySdev(-5) -0.00161 0.00023 -6.979 0
- Log Likelihood = -1816.05
- Schwarz Criterion = -1839.13
- Hannan-Quinn Criterion = -1831.35
- Akaike Criterion = -1826.05
- Sum of Squares = 2.45312e+016
- R-Squared = 0.6068
- R-Bar-Squared = 0.5679
- Residual SD = 1.64178e+007
- Residual Skewness = 0.1961
- Residual Kurtosis = 2.939
- Jarque-Bera Test = 0.6628 {0.718}
- Box-Pierce (residuals): Q(12) = 53.5449 {0}
- Box-Pierce (squared residuals): Q(12) = 29.3122 {0.004}
- Cointegration Test (ADF,0 lags) = .NaN {?}
- Cointegration Test (PP) = -5.4379 {?}*
- Lag length for ADF test selected by Akaike criterion.
- Covariance matrix from robust formula.
- * PP bandwidth = 6 (Parzen kernel.)
- ...Run completed in 0.44


1.26e+009
1.24e+009
1.22e+009
1.2e+009
1.18e+009
1.16e+009
1.14e+009
1.12e+009
Actual JPYUSDbillion
Fitted
20
40
60
80
100

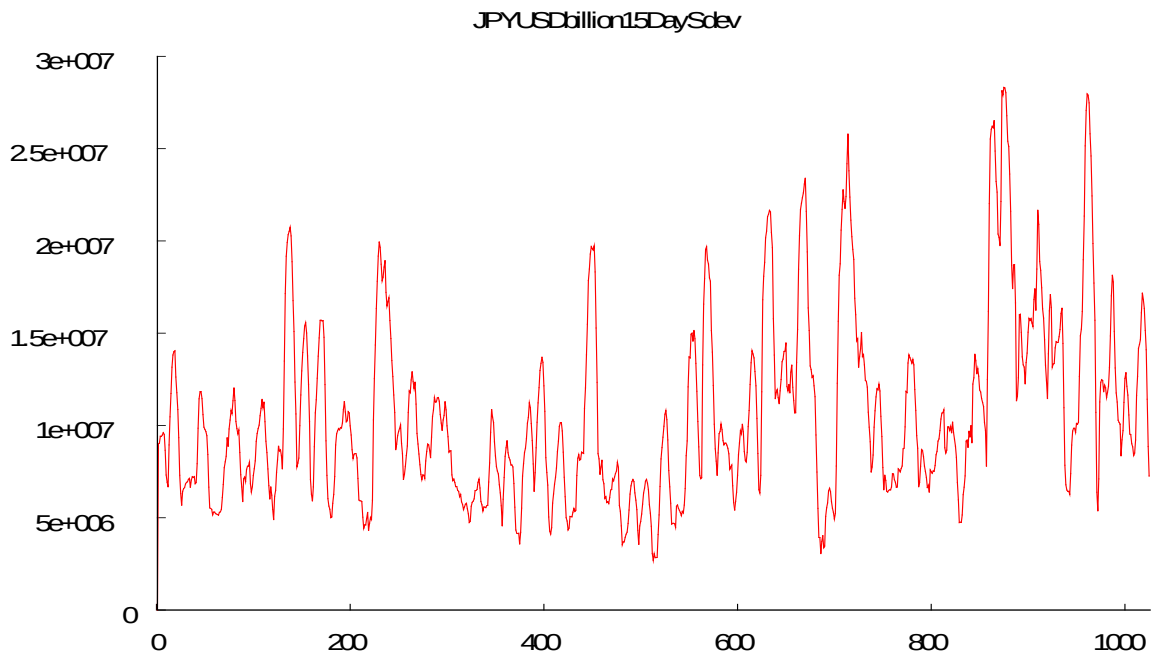
4e+007
3e+007
2e+007
1e+007
0
-1e+007
-2e+007
-3e+007
-4e+007
-5e+007
Residuals
2SE Bands
20
40
60
80
100

1-Period forecasts based on above regression model for 50 periods:



A striking improvement over basic I-Flow model for the first few periods, though not adding any more information than the simple AR(1) system from before. This improvement in forecasts seen after the introduction of the transformed volatility of the I-Flow variables, takes us to the next stage of modeling where we'll try and design a model to forecast the 15-Day moving standard deviation series of the JPYUSD pair.

Initial look at the 15-Day moving standard deviation series of the JPYUSD pair:



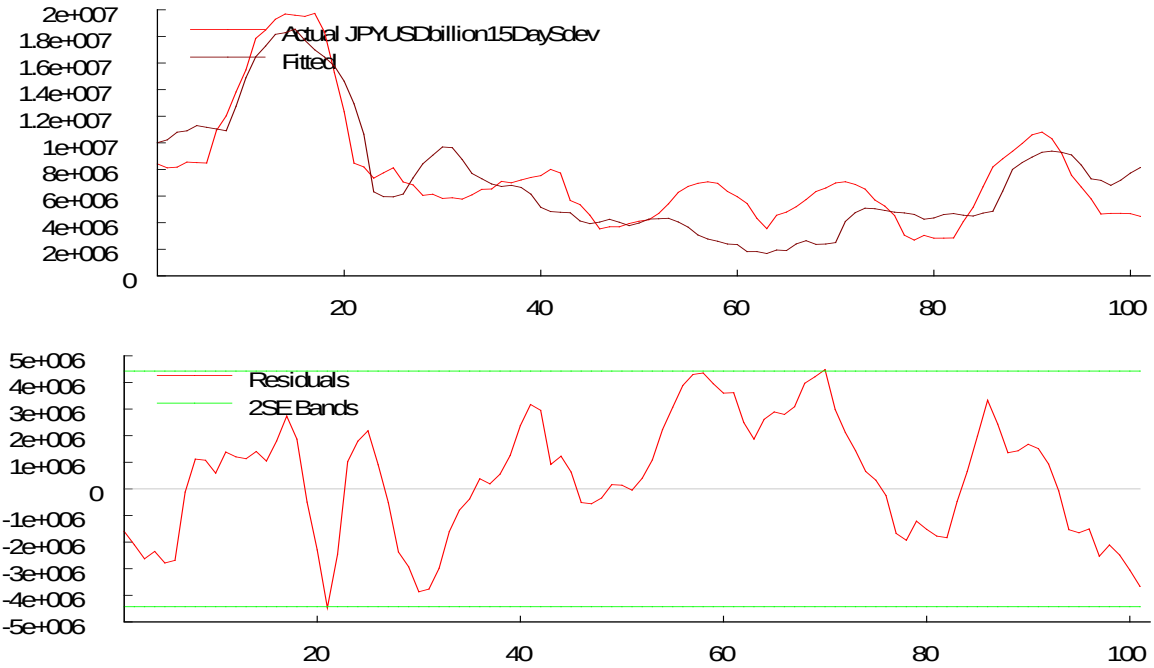
From above we conclude that the series is a stationary series exhibiting consistent periodic swings. We now proceed to test for a system utilizing the transformed I-Flow series tracking the above observed "volatility swing" series.

We design the volatility swing forecasting model on lines similar to the price series forecasting model by progressively dropping insignificant variables based on P-Values.

Initial Tests for five-period forecasts: Period 450 – 550

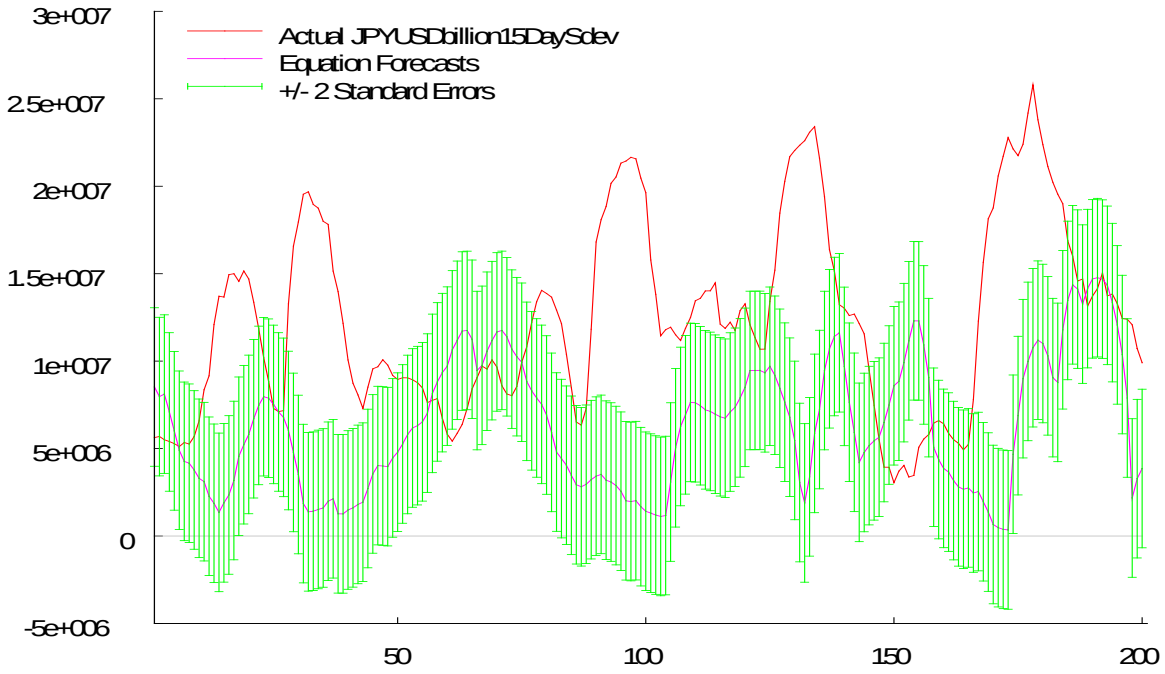
Stable model designed after progressively dropping insignificant variables based on P-Values.

- **********
- TSM4.29.22-01-09 Run 153 at 13:13:01 on 7-06-2009
- Data file is C:\Users\quantabaye\Documents\JPYUSDFlowV1.csv
- ---
- Dependent Variable is JPYUSDbillion15DaySdev
- 101 observations (450-550) used for estimation
- with 449 pre-sample observations.
- Estimation Method: Ordinary Least Squares
    - Estimate Std. Err. t Ratio p-Value
- JPFI:5-10YearsCumulative15DaySdev(-5) -0.00015 4e-005 -3.635 0
- JPFI:1-2YearsCumulative15DaySdev(-5) 0.00117 0.00011 10.665 0
- JPEQ/FITotalCumulative15DaySdev(-5) 0.00011 2e-005 5.437 0
    - Log Likelihood = -1619.81
    - Schwarz Criterion = -1626.73
- Hannan-Quinn Criterion = -1624.4
    - Akaike Criterion = -1622.81
    - Sum of Squares = 5.0352e+014
        - R-Squared = 0.7506
    - R-Bar-Squared = 0.7455
        - Residual SD = 2.22502e+006
    - Residual Skewness = 0.4429
    - Residual Kurtosis = 2.2679
    - Jarque-Bera Test = 5.5582 {0.062}
- Box-Pierce (residuals): Q(12) = 165.319 {0}
- Box-Pierce (squared residuals): Q(12) = 87.385 {0}
- Covariance matrix from robust formula.
- ...Run completed in 0.26


2e+007
1.8e+007
1.6e+007
1.4e+007
1.2e+007
1e+007
8e+006
6e+006
4e+006
2e+006
0
Actual JPYUSDbillion15DaySdev
Fitted
20
40
60
80
100

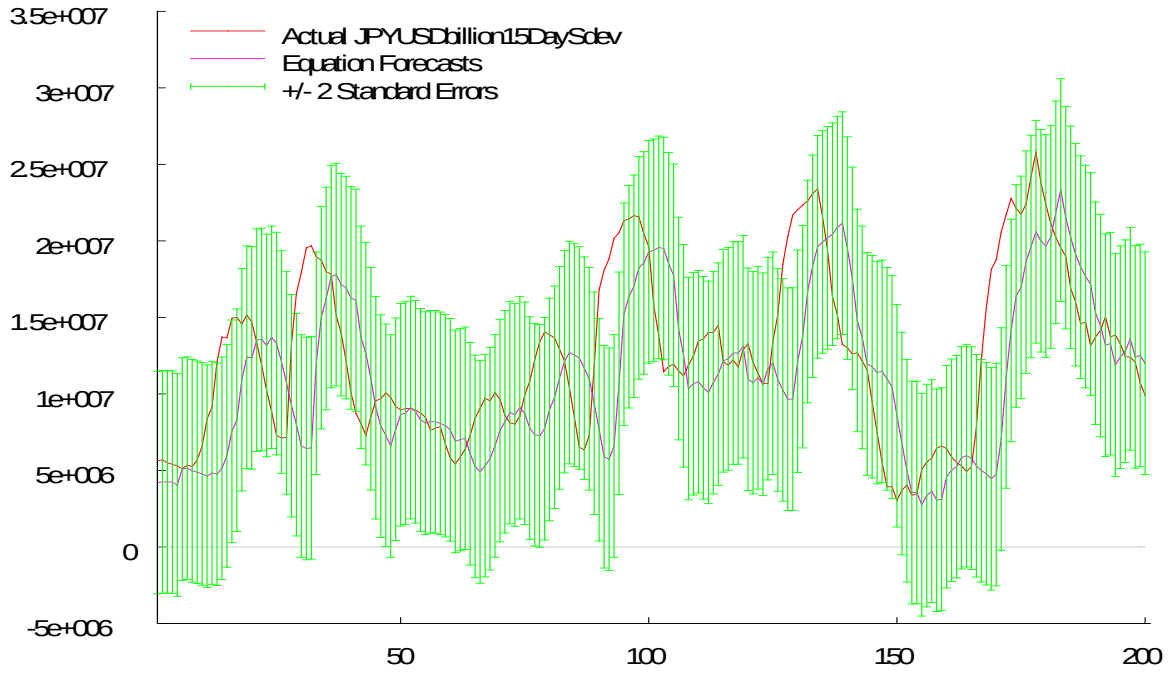
5e+006
4e+006
3e+006
2e+006
1e+006
0
-1e+006
-2e+006
-3e+006
-4e+006
-5e+006
Residuals
2SE Bands
20
40
60
80
100

5-Period forecasts based on above regression model for 200-Periods: Model run with I-Flow variables entering as 5-period lags, hence at any point in time the model returns forecasts for next 5 periods.



The above forecasts throwing up a very interesting pattern of Actual volatility swing series crossing forecasts from below and going on to take large swings consistently for very long periods outside sample (200 days).

Next we check the forecasting performance of an AR(5) forecast system (same series entering the model exclusive with a similar 5-day lag) for the series.

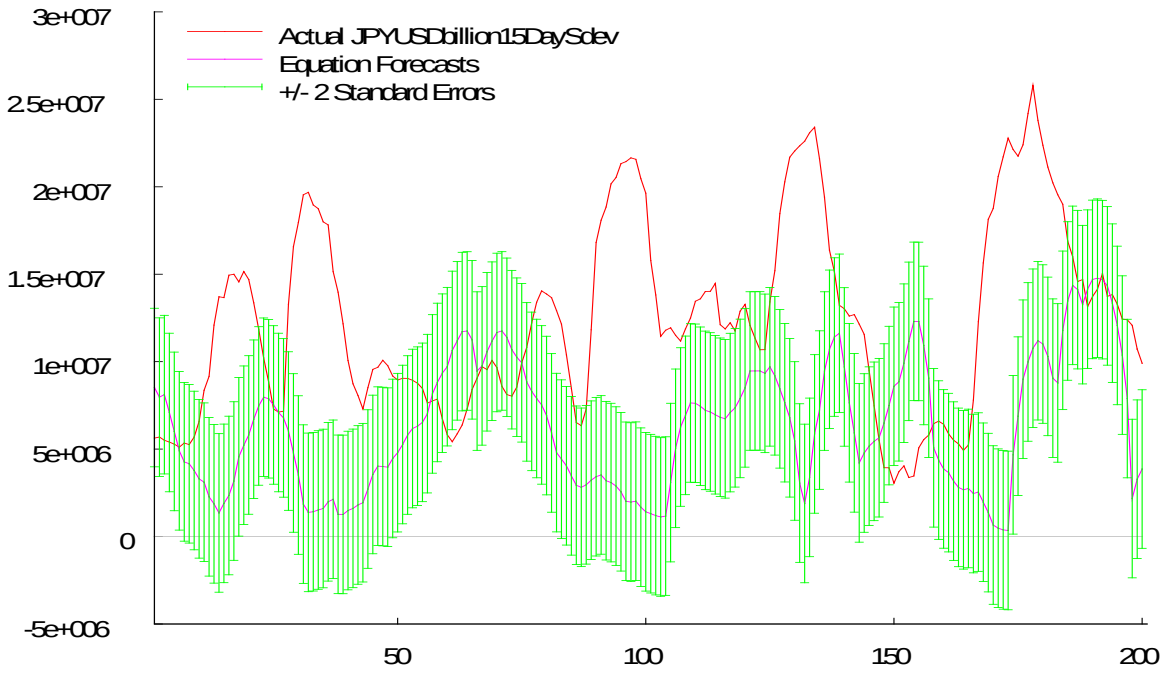


From above, it can be noticed that that the AR(5) system tracks the volatility swing series very closely and offers a similar pattern of crossings as the I-Flow model. But the striking contrast of the I-Flow model to the simple AR(5) is the fact that the I-Flow model crossings occur before "large volatility swings" that peak well above the other small peaks. The "I-Flow model", from this perspective, seems to offer a system of tracking large volatility swings and ignoring smaller not-so-significant ones.

Proposed Model 1:

The 5-Period forecasting model from above could be employed to time FX options trades (volatility bets). The initial crossing of the actual series of the forecast series from below can be caught well ahead of an impending large and pronounced volatility upswing.

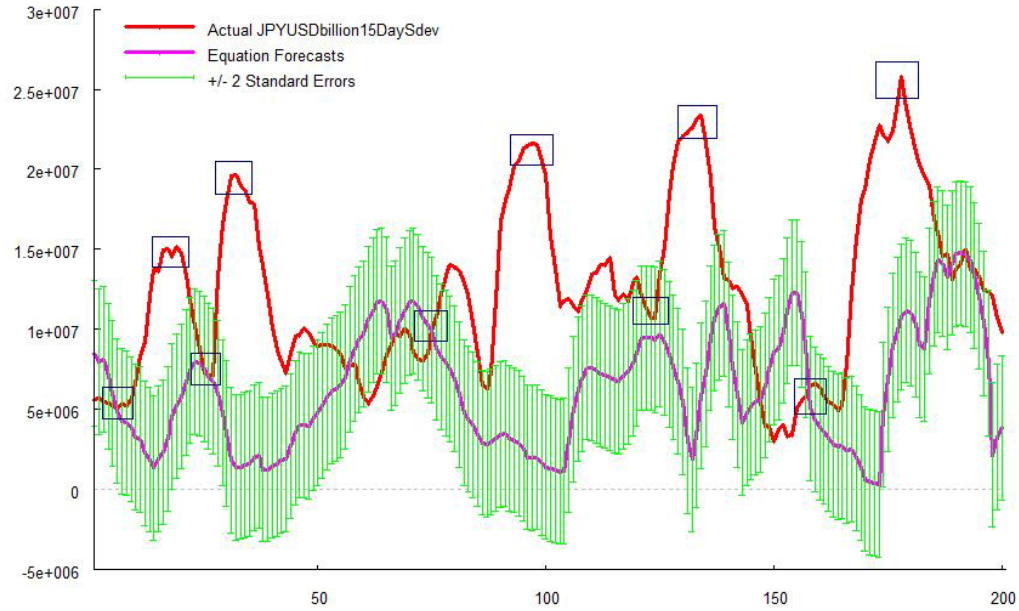
Next I overlay the above forecasts over the price series for the same period for the JPYUSD pair to check for informational advantage, if any, that the swing volatility forecasting system could offer in profitable long/short bets on the JPYUSD price series' trends.

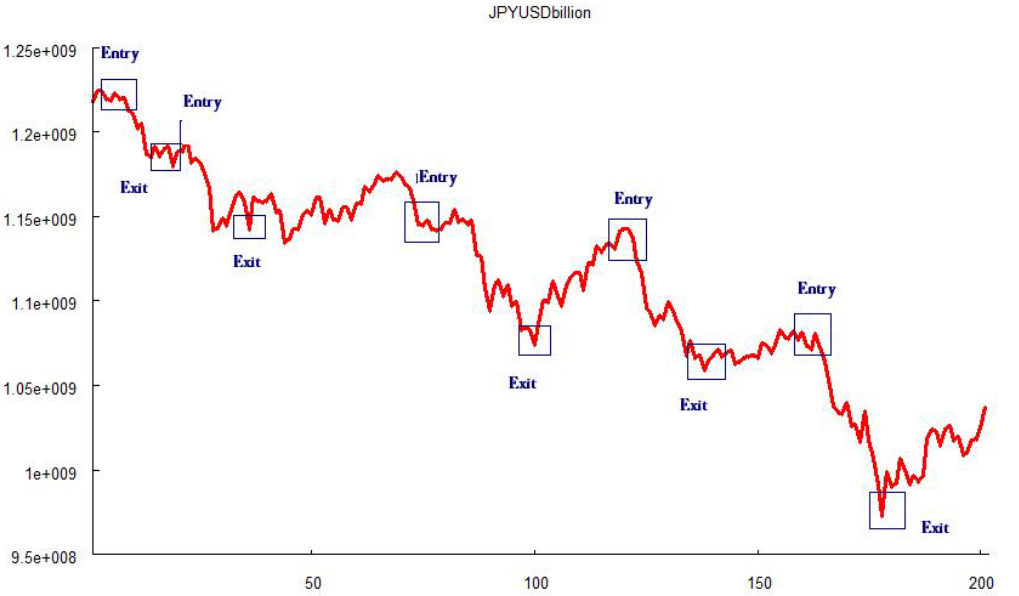


Proposed Model 2:

Above "Overlay System" points to very interesting patterns of the volatility swing model helping in timing of trends in the price series. The points of crossings of the actual swing series and its forecasts seem to help in timing the start of a trending period (Downward trend in the above example) and the subsequent "Peaking" of the respective upward swings seem to clearly help catch the other end of the previously identified trend.

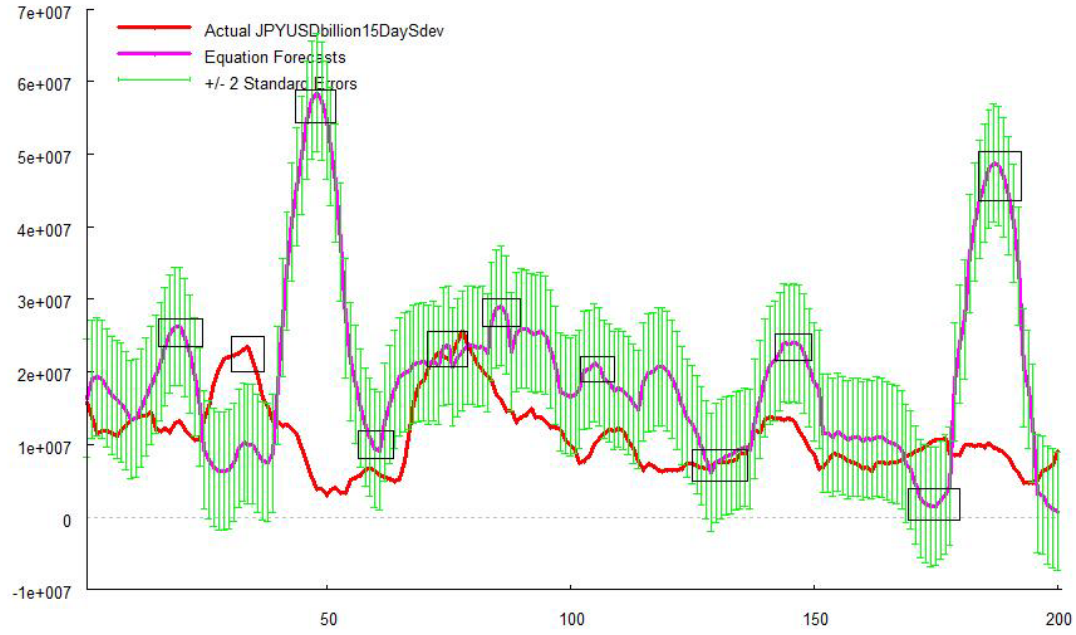
The figure below represents the above argument wherein the downtrend in the JPYUSD price series, if recognized by, for example, a technical trading system, the timing of "Entry" & "Exit" points for such a system could be timed with very fine precision.



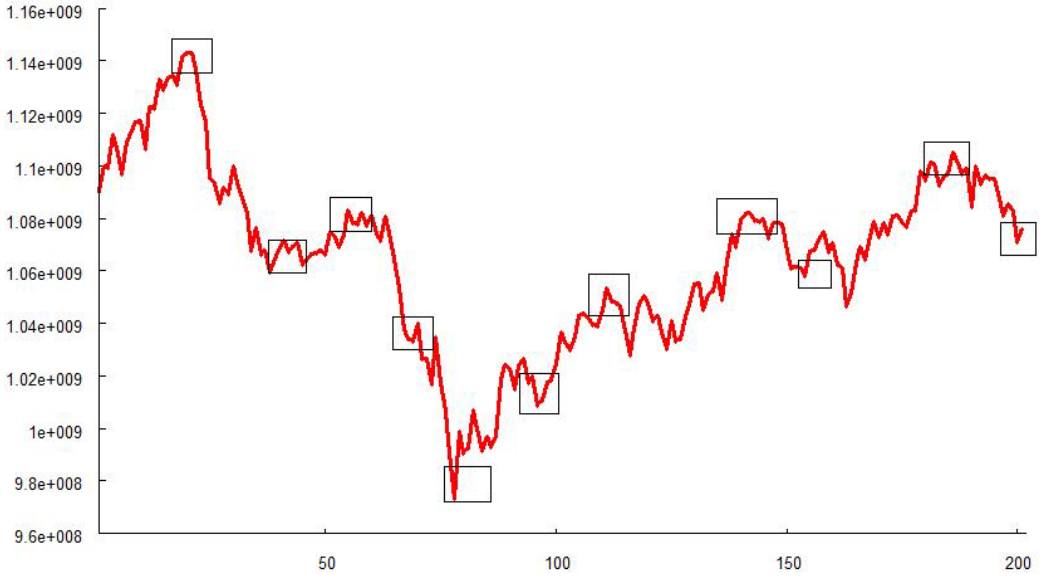


Same model run for periods 550-650 and a 200 period 5-Day forecasts system.

It is clearly evident that even during overlap periods (the first 100 period forecasts for this model are same as the last 100 periods of the above model) peaks and crossings offer clear “Entry” & “Exit” points that match the corresponding points from above for the overlap period. This provides a very strong evidence for the stability of the “Volatility Swing Forecasting I-Flow system” designed and explored in this paper.



JPYUSDbillion



Conclusion: Cross-currency flow data maintained by large institutions (custodian banks) seem to, on initial examination, hold information that has potential to provide informational content to a very intelligent and complex statistical trading system for catching "Entry" &"Exit" points for a traditional trend prediction system and also a stand-alone model for volatility bets involving FX options trades. This initial investigation points to the design of an intelligent trading system that could add value to FX trades of 1-day latency.

Notes:

- JPYUSD data was scaled up by a factor of a billion to accommodate the series to the natural scale of most of the I-Flow data series.
- The I-Flow series were subjected to a I(1) Vs I(0) initial non-stationarity tests (P-P & KPSS tests) before the choice of the series to be considered for the model was made. His because the JPYUSD series, as expected, was found to be I(1) itself.