News - Good or Bad - and Its Impact On Volatility Predictions over Multiple Horizons

Authors: Xilong Chen  Eric Ghysels

January 24, 2010 Version
Outline

1. Introduction

2. Volatility Measurement and Model Specifications
   - Measuring Volatility
   - Model Specifications

3. Empirical Results
   - Is News Impact Asymmetric?
   - Out-of-sample Forecasting Performance
   - Are the Forecasting Gains Statistically Significant?
Motivation: Examine whether the sign and magnitude of intro-daily returns have impact on expected volatility the next day or over the longer future horizons.

Contributions:
- HAR-S-RV-J: introduce a new model that includes asymmetries in terms of semi-variances.
- Introduce a new class of semi-parametric/parametric models applicable to a mixture of high and low frequency data which feature asymmetries.
- Asymmetry matters for volatility forecasting.
- Models using intra-daily returns outperform models with daily realized variance and semi-variance measures in terms of out-of-sample forecasting.
Motivation: Examine whether the sign and magnitude of intra-daily returns have impact on expected volatility the next day or over the longer future horizons.

Contributions:

- HAR-S-RV-J: introduce a new model that includes asymmetries in terms of semi-variances.
- Introduce a new class of semi-parametric/parametric models applicable to a mixture of high and low frequency data which feature asymmetries.
- Asymmetry matters for volatility forecasting.
- Models using intra-daily returns outperform models with daily realized variance and semi-variance measures in terms of out-of-sample forecasting.
Main Results

- Moderate good (intra-daily) news reduce volatility (the next day), while both very good news (unusual high intra-daily positive returns) and bad news (negative returns) increase volatility, with the latter having a more severe impact. The asymmetries disappear over longer horizons.

Outline

1. Introduction

2. Volatility Measurement and Model Specifications
   - Measuring Volatility
   - Model Specifications

3. Empirical Results
   - Is News Impact Asymmetric?
   - Out-of-sample Forecasting Performance
   - Are the Forecasting Gains Statistically Significant?
Measuring Volatility

- **Realized Variance (RV):**

  \[ RV_{t,t+k}(\Delta^{-1}) \equiv \sum_{j=t}^{t+k} \sum_{i=1}^{\Delta^{-1}} r_{j,i\Delta}^2 \]

  \( r_{j,i\Delta} \) : the log asset price difference (return) over short time interval \( i \) of length \( \Delta \) on day \( j \).

- **Bi-power Variation (BPV):**

  \[ BPV_{t,t+k}(\Delta^{-1}) \equiv \sum_{j=t}^{t+k} \sum_{i=2}^{\Delta^{-1}} |r_{j,i\Delta}||r_{j,(i-1)\Delta}| \]
Measuring Volatility

- Realized Variance (RV):

\[ RV_{t,t+k}(\Delta^{-1}) \equiv \sum_{j=t}^{t+k} \Delta^{-1} \sum_{i=1}^{\Delta-1} r_{j,i}^2 \Delta \]

- Bi-power Variation (BPV):

\[ BPV_{t,t+k}(\Delta^{-1}) \equiv \sum_{j=t}^{t+k} \Delta^{-1} \sum_{i=2}^{\Delta-1} |r_{j,i}\Delta| |r_{j,(i-1)\Delta}| \]

\[ r_{j,i}\Delta \]: the log asset price difference (return) over short time interval \( i \) of length \( \Delta \) on day \( j \).
Measuring Volatility

Semi-variance:

\[
\text{Semi}V^{+}_{t+k,t} \equiv \sum_{j=t}^{t+k} \sum_{i=1}^{\Delta^{-1}} r_{j,i\Delta}^2 1_{r_{j,i\Delta} > 0}
\]

\[
\text{Semi}V^{-}_{t+k,t} \equiv RV_{t+k,t} - \text{Semi}V^{+}_{t+k,t}
\]
Outline

1. Introduction

2. Volatility Measurement and Model Specifications
   - Measuring Volatility
   - Model Specifications

3. Empirical Results
   - Is News Impact Asymmetric?
   - Out-of-sample Forecasting Performance
   - Are the Forecasting Gains Statistically Significant?
Models with Daily Volatility

- **Generic Equation:**

\[RV_{t,t+1} = \psi_0 + \sum_{j=0}^{\tau} \psi_j(\theta) RV_{t-j-1,t-j} + \epsilon_{t,t+1}\]

- **HAR-RV (HAR: Heterogeneous Autoregressive):**

\[RV_{t,t+k} = \psi_0 + \psi_D RV_{t-1,t} + \psi_W RV_{t-5,t} + \psi_M RV_{t-22,t} + \epsilon_{t,t+k}\]
Models with Daily Volatility

- **Generic Equation:**
  \[
  RV_{t,t+1} = \psi_0 + \sum_{j=0}^{\tau} \psi_j(\theta) RV_{t-j-1,t-j} + \epsilon_{t,t+1}
  \]

- **HAR-RV (HAR: Heterogeneous Autoregressive):**
  \[
  RV_{t,t+k} = \psi_0 + \psi_D RV_{t-1,t} + \psi_W RV_{t-5,t} + \psi_M RV_{t-22,t} + \epsilon_{t,t+k}
  \]
HAR-RV-J (J: Jump):

\[ RV_{t,t+k} = \psi_0 + \psi_D BPV_{t-1,t} + \psi_W BPV_{t-5,t} + \psi_M BPV_{t-22,t} + \psi_J J_t + \epsilon_{t,t+k} \]

HAR-S-RV-J (S: Semi):

\[ RV_{t,t+k} = \psi_0 + \psi_D^{+} SemiV^+_{t-1,t} + \psi_D^{-} SemiV^-_{t-1,t} + \psi_W^{+} SemiV^+_{t-5,t} + \psi_W^{-} SemiV^-_{t-5,t} + \psi_M^{+} SemiV^+_{t-22,t} + \psi_M^{-} SemiV^-_{t-22,t} + \psi_J J_t + \epsilon_{t,t+k} \]
HAR-RV-J (J: Jump):

\[ RV_{t,t+k} = \psi_0 + \psi_D BPV_{t-1,t} + \psi_W BPV_{t-5,t} \]
\[ + \psi_M BPV_{t-22,t} + \psi_J J_t + \epsilon_{t,t+k} \]

HAR-S-RV-J (S: Semi):

\[ RV_{t,t+k} = \psi_0 + \psi^+_D SemiV^+_{t-1,t} + \psi^-_D SemiV^-_{t-1,t} \]
\[ + \psi^+_W SemiV^+_{t-5,t} + \psi^-_W SemiV^-_{t-5,t} \]
\[ + \psi^+_M SemiV^+_{t-22,t} + \psi^-_M SemiV^-_{t-22,t} \]
\[ + \psi_J J_t + \epsilon_{t,t+k} \]
**Models with Intro-daily Returns: Semi-parametric**

- **SP: Semi-parametric MIDAS regression**

\[ RV_{t,t+1} = \psi_0 + \sum_{j=1}^{\tau} \sum_{i=1}^{\Delta^{-1}} \psi_{ij}(\theta) \text{NIC}(r_{t-j,i\Delta}) + \epsilon_{t,t+1} \]

- **MIDAS: Mixed Data Sampling**
- \( \psi_{ij}(\theta) \): polynomial lag structure parameterized by \( \theta \)
- \( \text{NIC}(.) \): the news impact curve
- Semi-parametric: a parametric estimation of \( \psi_{ij}(\theta) \) and a non-parametric \( \text{NIC}(.) \)
- SP model nests the generic equation when setting \( \psi_{ij} = \psi_i \forall j = 1, \ldots, \Delta^{-1} \), and \( \text{NIC}(r) \equiv r^2 \)
Models with Intro-daily Returns: Parametric

- **SYMM:** Symmetric NIC
  \[ NIC(r) = br^2 \]

- **ASYMGJR:** Asymmetric GJR
  \[ NIC(r) = br^2 + cr^21_{r<0} \]
  GJR Model: Glosten, Jagannathan, and Runkle (1993)

- **ASYML:** Asymmetric Location Shift
  \[ NIC(r) = b(r - c)^2 \]
Models with Intro-daily Returns: Parametric

- **SYMM**: Symmetric NIC
  \[ NIC(r) = br^2 \]

- **ASYMGJR**: Asymmetric GJR
  \[ NIC(r) = br^2 + cr^2 1_{r<0} \]

  GJR Model: Glosten, Jagannathan, and Runkle (1993)

- **ASYMLS**: Asymmetric Location Shift
  \[ NIC(r) = b(r - c)^2 \]
Models with Intro-daily Returns: Parametric

- **SYMM: Symmetric NIC**
  \[ NIC(r) = br^2 \]

- **ASYMGJR: Asymmetric GJR**
  \[ NIC(r) = br^2 + cr^2 1_{r<0} \]
  GJR Model: Glosten, Jagannathan, and Runkle(1993)

- **ASYMLS: Asymmetric Location Shift**
  \[ NIC(r) = b(r - c)^2 \]
Polynomial lag structure $\psi_{ij}(\theta)$

- To Nest models based on $RV$ and SP

\[ RV_{t,t+1} = \psi_0 + \sum_{j=0}^{\tau} \psi_j(\theta) RV_{t-j-1,t-j} + \epsilon_{t,t+1} \]

\[ RV_{t,t+1} = \psi_0 + \sum_{j=1}^{\tau} \sum_{i=1}^{\Delta^{-1}} \psi_{ij}(\theta) NIC(r_{t-j,i\Delta}) + \epsilon_{t,t+1} \]
Polynomial lag structure $\psi_{ij}(\theta)$ (cont.)

Let

$$\psi_{ij}(\theta) \equiv \psi^D_j(\theta) \times \psi^{ID}_i(\theta)$$

$\psi^D_j(\theta)$: a daily weighting scheme
$\psi^{ID}_i(\theta)$: the intra-daily weights

With equal intro-daily weights and quadratic NIC, the above two models are the same

the authors adopt the following specification of polynomials

$$\psi^D_j(\theta) \times \psi^{ID}_i(\theta) = Beta(j, \tau, \theta_1, \theta_2) \times Beta(i, \Delta^{-1}, \theta_3, \theta_4)$$
Estimation Issues

1. Estimation of parametric models
   - OLS for HAR-RV, HAR-RV-J, and HAR-S-RV-J
   - Nonlinear least squares for models those use intra-daily returns, SYMM, ASYMGJR, and ASYMLS

2. Estimation of the semi-parametric MIDAS regression model and non parametric estimation of the NIC function

3. Data

<table>
<thead>
<tr>
<th></th>
<th>Cash</th>
<th>Futures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dow Jones</td>
<td>DJC</td>
<td>DJF</td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>SPC</td>
<td>SPF</td>
</tr>
</tbody>
</table>
Is News Impact Asymmetric using Semi-variances?

- **HAR-S-RV-J model**
- Entries are p-values
- The first three columns test $\psi_D^+ = \psi_D^-$, $\psi_W^+ = \psi_W^-$, $\psi_M^+ = \psi_M^-$, respectively. The final column reports joint tests.
- The daily lag coefficient do not seem to feature asymmetries for predicting future daily volatility, yet it does for weekly and monthly forecasts.
- The asymmetries are present at all forecast horizons based on the final column, except for DJC.
### Is News Impact Asymmetric using Intro-daily Returns?

<table>
<thead>
<tr>
<th></th>
<th>( \theta_1 )</th>
<th>( \theta_2 )</th>
<th>( \theta_3 )</th>
<th>( \theta_4 )</th>
<th>( a )</th>
<th>( b )</th>
<th>( c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Futures Market - One-day horizon</td>
<td>( 2.337e-014 )</td>
<td>3.051</td>
<td>1.360</td>
<td>1.955</td>
<td>-0.008254</td>
<td>-125.9</td>
<td>428.5</td>
</tr>
<tr>
<td>(std. dev.)</td>
<td>(0.005456)</td>
<td>(0.03413)</td>
<td>(0.001195)</td>
<td>(0.004495)</td>
<td>(0.01655)</td>
<td>(4.728)</td>
<td>(10.08)</td>
</tr>
<tr>
<td>ASYMLS</td>
<td>( 2.337e-014 )</td>
<td>1.974</td>
<td>1.239</td>
<td>2.825</td>
<td>-39.18</td>
<td>67.61</td>
<td>0.7621</td>
</tr>
<tr>
<td>(std. dev.)</td>
<td>(0.004674)</td>
<td>(0.0215)</td>
<td>(0.002814)</td>
<td>(0.01108)</td>
<td>(0.4254)</td>
<td>(0.7221)</td>
<td>(0.00199)</td>
</tr>
</tbody>
</table>

\[ RV_{t,t+1} = \psi_0 + \sum_{j=1}^{\tau} \sum_{i=1}^{\Delta^{-1}} \psi_{ij}(\theta)NIC(r_{t-j,i\Delta}) + \epsilon_{t,t+1} \]

\[ \psi_{ij}(\theta) = Beta(j, \tau, \theta_1, \theta_2) \times Beta(i, \Delta^{-1}, \theta_3, \theta_4) \]

**ASYMGJR:**

\[ NIC(r) = br^2 + cr^21_{r<0} \]

**ASYMLS:** \[ NIC(r) = b(r - c)^2 \]

The intro-daily asymmetries is present based on the significance of parameter \( c \).
Is News Impact Asymmetric? (cont.)

- SPF news impact curve via the SP estimation.
- X-axis: annualized intra-daily returns.
- Y-axis: Impact on next day, week, and month’s $RV$ scaled by the mean of $RV$ for each horizon.
- Daily and weekly NIC: asymmetry. Monthly NIC: symmetric. Bad news has more acute impact than positive news.
Outline

1. Introduction

2. Volatility Measurement and Model Specifications
   - Measuring Volatility
   - Model Specifications

3. Empirical Results
   - Is News Impact Asymmetric?
   - Out-of-sample Forecasting Performance
   - Are the Forecasting Gains Statistically Significant?
## Out-of-sample Forecasting Performance

<table>
<thead>
<tr>
<th>Period</th>
<th>One-day ahead forecasts</th>
<th>One-week ahead forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DJC</td>
<td>DJF</td>
</tr>
<tr>
<td><strong>SP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005-2006</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>SYMM</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005-2006</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2007-2008</td>
<td>0.95</td>
<td>1.10</td>
</tr>
<tr>
<td><strong>ASYMGJR</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005-2006</td>
<td>0.75</td>
<td>0.60</td>
</tr>
<tr>
<td>2007-2008</td>
<td>0.95</td>
<td>1.41</td>
</tr>
<tr>
<td><strong>ASYMLS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005-2006</td>
<td>1.25</td>
<td>1.20</td>
</tr>
<tr>
<td>2007-2008</td>
<td>0.91</td>
<td>1.02</td>
</tr>
</tbody>
</table>

### Parametric Intra-daily returns

- **SYMM**
  - 2005-2006: 1.00, 1.00, 1.00, 0.57, 1.50, 1.40, 0.67, 0.25
  - 2007-2008: 0.95, 1.10, 1.79, 2.88, 0.79, 1.19, 0.60, 3.67

- **ASYMGJR**
  - 2005-2006: 0.75, 0.60, 0.75, 0.72, 1.00, 1.40, 0.67, 0.25
  - 2007-2008: 0.95, 1.41, 1.11, 1.29, 0.67, 1.11, 0.42, 6.72

- **ASYMLS**
  - 2005-2006: 1.25, 1.20, 0.57, 1.29, 1.50, 1.40, 0.67, 0.38
  - 2007-2008: 0.91, 1.02, 1.01, 1.12, 0.73, 0.95, 0.48, 3.67

### Parametric Daily Volatility Measures

- **HAR-RV**
  - 2005-2006: 1.25, 1.00, 1.75, 0.57, 3.00, 1.60, 3.00, 0.63
  - 2007-2008: 1.04, 1.13, 1.08, 1.17, 1.06, 1.17, 0.59, 1.00

- **HAR-RV-J**
  - 2005-2006: 1.00, 0.80, 0.86, 2.00, 1.20, 0.80, 2.33, 0.63
  - 2007-2008: 1.02, 1.09, 1.12, 1.22, 1.02, 0.85, 0.60, 0.99

- **HAR-S-RV-J**
  - 2005-2006: 1.00, 0.60, 1.00, 0.43, 1.50, 0.40, 1.00, 0.50
  - 2007-2008: 1.08, 0.98, 1.02, 1.04, 1.03, 0.84, 0.53, 0.97

- **MSFE**

\[
MSFE_k^m = \frac{1}{N_k} \sum_{i=1}^{N_k} (u_{t+(i-1)k,t+ik})^2
\]

\[
u_{t,t+k} = RV_{t,t+k} - \hat{RV}_{t,t+k}
\]

### Out-of-sample MSFE of the SP model is the benchmark. Values below one outperform the benchmark.

### The level of MSFE are different.

### The best models always feature asymmetries, but not the same model.
Outline

1. Introduction

2. Volatility Measurement and Model Specifications
   - Measuring Volatility
   - Model Specifications

3. Empirical Results
   - Is News Impact Asymmetric?
   - Out-of-sample Forecasting Performance
   - Are the Forecasting Gains Statistically Significant?
### Are the Forecasting Gains Statistically Significant?

<table>
<thead>
<tr>
<th></th>
<th>One-day ahead forecasts</th>
<th>One-week ahead forecasts</th>
<th>One-month ahead forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DJC</td>
<td>DJF</td>
<td>SPC</td>
</tr>
<tr>
<td>SYMM</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ASYMGJR</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>ASYMLS</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Evaluation sample: 2001-2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYMM</td>
<td>0.03 0.00 0.11 0.00 0.01 0.06 0.02 0.06 0.09 0.15 0.11 0.21</td>
</tr>
<tr>
<td>ASYMGJR</td>
<td>0.03 0.06 0.03 0.04 0.01 0.02 0.02 0.06 0.15 0.13 0.11 0.25</td>
</tr>
<tr>
<td>ASYMLS</td>
<td>0.02 0.00 0.04 0.09 0.02 0.07 0.03 0.06 0.11 0.14 0.12 0.26</td>
</tr>
</tbody>
</table>

Giacomini and White (2006) test: \( H_0 : E[(u_{t+k}^b)^2 - (u_{t+k}^m)^2]|I_{t-1}] = 0 \)

- \( b \): HAR-S-RV-J Benchmark; \( m \): model \( m \); \( I_{t-1} \): information set(constant and \((u_{t-k}^b)^2 \) here); the test statistic will be \( \chi^2_1 \)
- All entries are p-values. Boldfaced: the benchmark model id better than the alternative model.
- For one-day ahead forecasts: ASYMGJR outperform the benchmark for both periods; SYMM is inferior. Adding the crisis, the evidence weakens the evidence against ASYMLS.
Summary

1. New impact curve is asymmetric for short term horizons and the asymmetries disappear over longer horizons.
