Asymmetric Effects and Long Memory in the Volatility of DJIA Stocks

Marcel Scharth and Marcelo C. Medeiros

Pontifical Catholic University of Rio de Janeiro

November 7, 2006
Introduction and Motivation

Some stylized facts of financial time series

- The distribution of financial returns has fat tails: Kurtosis much larger than 3.
- Volatility clustering.
- The volatility of returns displays long-range dependence: Autocorrelations decay very slowly.
- Asymmetries

Modeling the dynamics of financial time series

- LATENT VOLATILITY MODELS: GARCH FAMILY, STOCHASTIC VOLATILITY, EWMA, ETC.
- ESTIMATE VOLATILITY WITH INTRADAY DATA (REALIZED VOLATILITY) AND USE STANDARD TIME SERIES TECHNIQUES – SOLUTION ADOPTED IN THIS PAPER!
- IMPORTANT FOR RISK MANAGEMENT AND ASSET PRICING!

Marcel Scharth and Marcelo C. Medeiros
Asymmetries and Long Memory in Volatility
Introduction and Motivation

Some stylized facts of financial time series

- The distribution of financial returns has fat tails: Kurtosis much larger than 3.
- Volatility clustering.
- The volatility of returns displays long-range dependence: Autocorrelations decay very slowly.
- Asymmetries

Modeling the dynamics of financial time series

- 
- 
- 

Marcel Scharth and Marcelo C. Medeiros

Asymmetries and Long Memory in Volatility
Introduction and Motivation

Some stylized facts of financial time series

- The distribution of financial returns has fat tails: Kurtosis much larger than 3.
- Volatility clustering.
- The volatility of returns displays long-range dependence: Autocorrelations decay very slowly.
- Asymmetries

Modeling the dynamics of financial time series

- Latent volatility models: GARCH, stochastic volatility, EWMA, etc.
Introduction and Motivation

Some stylized facts of financial time series

- The distribution of financial returns has fat tails: Kurtosis much larger than 3.
- Volatility clustering.
- The volatility of returns displays long-range dependence: Autocorrelations decay very slowly.
- Asymmetries

Modeling the dynamics of financial time series

- Latent volatility models: GARCH family, stochastic volatility, EWMA, etc.
- Estimate volatility with intraday data (realized volatility) and use standard time series techniques – Solution adopted in this paper!
Introduction and Motivation

Some stylized facts of financial time series

- The distribution of financial returns has fat tails: Kurtosis much larger than 3.
- Volatility clustering.
- The volatility of returns displays long-range dependence: Autocorrelations decay very slowly.
- Asymmetries

Modeling the dynamics of financial time series

- Latent volatility models: GARCH family, stochastic volatility, EWMA, etc.
- Estimate volatility with intraday data (realized volatility) and use standard time series techniques – Solution adopted in this paper!
- Important for risk management and asset pricing!
Some stylized facts of financial time series

- The distribution of financial returns has fat tails: Kurtosis much larger than 3.
- Volatility clustering.
- The volatility of returns displays long-range dependence: Autocorrelations decay very slowly.
- Asymmetries

Modeling the dynamics of financial time series

- Latent volatility models: GARCH family, stochastic volatility, EWMA, etc.
- Estimate volatility with intraday data (realized volatility) and use standard time series techniques – Solution adopted in this paper!
- Important for risk management and asset pricing!
Introduction and Motivation

Some stylized facts of financial time series

- The distribution of financial returns has fat tails: Kurtosis much larger than 3.
- Volatility clustering.
- The volatility of returns displays long-range dependence: Autocorrelations decay very slowly.
- Asymmetries

Modeling the dynamics of financial time series

- Latent volatility models: GARCH family, stochastic volatility, EWMA, etc.
- Estimate volatility with intraday data (realized volatility) and use standard time series techniques – Solution adopted in this paper!
- Important for risk management and asset pricing!
Introduction and Motivation

Some stylized facts of financial time series

- The distribution of financial returns has fat tails: Kurtosis much larger than 3.
- Volatility clustering.
- The volatility of returns displays long-range dependence: Autocorrelations decay very slowly.
- Asymmetries

Modeling the dynamics of financial time series

- Latent volatility models: GARCH family, stochastic volatility, EWMA, etc.
- Estimate volatility with intraday data (realized volatility) and use standard time series techniques – Solution adopted in this paper!
- Important for risk management and asset pricing!
Introduction and Motivation

Example


![Graph of daily returns of IBM from 1994 to 2003 with a histogram showing kurtosis of 7.6857]
Example

Introduction and Motivation

Example

Introduction and Motivation

Modeling long-range dependence

- Fractionally integrated (ARFIMA) models?
- Regime-switching?
- Structural breaks?

Modeling asymmetries
Introduction and Motivation

Modeling long-range dependence

- Fractionally integrated (ARFIMA) models?
- Regime-switching?
- Structural breaks?

Modeling asymmetries

- 
- 
- 

Marcel Scharth and Marcelo C. Medeiros

Asymmetries and Long Memory in Volatility
Introduction and Motivation

Modeling long-range dependence
- Fractionally integrated (ARFIMA) models?
- Regime-switching?
- Structural breaks?

Modeling asymmetries
- Regime-switching
- Structural breaks
Introduction and Motivation

Modeling long-range dependence
- Fractionally integrated (ARFIMA) models?
- Regime-switching?
- Structural breaks?

Modeling asymmetries
- Regime-switching
- How many regimes?
Introduction and Motivation

Modeling long-range dependence
- Fractionally integrated (ARFIMA) models?
- Regime-switching?
- Structural breaks?

Modeling asymmetries
- Regime-switching
- How many regimes?
Introduction and Motivation

Modeling long-range dependence
- Fractionally integrated (ARFIMA) models?
- Regime-switching?
- Structural breaks?

Modeling asymmetries
- Regime-switching
- How many regimes?
Introduction and Motivation

Modeling long-range dependence
- Fractionally integrated (ARFIMA) models?
- Regime-switching?
- Structural breaks?

Modeling asymmetries
- Regime-switching
- How many regimes?
Realized Volatility: Definition and Main Results

Definition

- Realized variance is defined as the sum of all available intraday high frequency squared returns.
- Realized volatility is the square root of the realized variance.
- Under the assumption of uncorrelated intraday returns, the realized variance is a consistent estimator of the integrated variance in a continuous-time diffusion model – Andersen et al. (Econometrica, 2003) and Barndorff-Nielsen and Shephard (JRSS-B, 2002).
Realized Volatility: Definition and Main Results

Definition

- Realized variance is defined as the sum of all available intraday high frequency squared returns.
- Realized volatility is the square root of the realized variance.
- Under the assumption of uncorrelated intraday returns, the realized variance is a consistent estimator of the integrated variance in a continuous-time diffusion model – Andersen et al. (Econometrica, 2003) and Barndorff-Nielsen and Shephard (JRRS-B, 2002).
Realized Volatility: Definition and Main Results

**Definition**

- Realized variance is defined as the sum of all available intraday high frequency squared returns.
- **Realized volatility** is the square root of the realized variance.
- Under the assumption of uncorrelated intraday returns, the realized variance is a consistent estimator of the integrated variance in a continuous-time diffusion model – Andersen et al. (Econometrica, 2003) and Barndorff-Nielsen and Shephard (JRRS-B, 2002).
Realized Volatility: Definition and Main Results

Definition

- Realized variance is defined as the sum of all available intraday high frequency squared returns.
- Realized volatility is the square root of the realized variance.
- Under the assumption of uncorrelated intraday returns, the realized variance is a consistent estimator of the integrated variance in a continuous-time diffusion model – Andersen et al. (Econometrica, 2003) and Barndorff-Nielsen and Shephard (JRSS-B, 2002).
Realized Volatility: Definition and Main Results

Problem
Due to microstructure noise, the intraday returns are autocorrelated. Thus, the realized variance is not a consistent estimator of the integrated variance.

Solution
- Use the two time scales estimator (TTSE) put forward by Zhang, Mykland, and Aït-Sahalia (JASA, 2005).
Problem

Due to microstructure noise, the intraday returns are autocorrelated. Thus, the realized variance is not a consistent estimator of the integrated variance.

Solution

- Use the two time scales estimator (TTSE) put forward by Zhang, Mykland, and Aït-Sahalia (JASA, 2005).
Main Contribution

Key Idea

- Use an unified tree-structured framework (model) to deal with structural breaks and regime-shifts.
- Combination of regression trees and smooth transition models.

Main Advantages

- Nests several nonlinear models previously proposed.
- Genuinely different regimes.
- Multiple transition variables.
- Long-range dependence and intermittent dynamics.
Main Contribution

Key Idea

- Use an unified tree-structured framework (model) to deal with structural breaks and regime-shifts.
- Combination of regression trees and smooth transition models.

Main Advantages

- Nests several nonlinear models previously proposed.
- Genuinely different regimes.
- Multiple transition variables.
- Long-range dependence and intermittent dynamics.
Main Contribution

Key Idea

- Use an unified tree-structured framework (model) to deal with structural breaks and regime-shifts.
- Combination of regression trees and smooth transition models.

Main Advantages

- Nests several nonlinear models previously proposed.
- Genuinely different regimes.
- Multiple transition variables.
- Long-range dependence and intermittent dynamics.
Main Contribution

Key Idea

- Use an unified tree-structured framework (model) to deal with structural breaks and regime-shifts.
- Combination of regression trees and smooth transition models.

Main Advantages

- Nests several nonlinear models previously proposed.
- Genuinely different regimes.
- Multiple transition variables.
- Long-range dependence and intermittent dynamics.
Main Contribution

Key Idea

- Use an unified tree-structured framework (model) to deal with structural breaks and regime-shifts.
- Combination of regression trees and smooth transition models.

Main Advantages

- Nests several nonlinear models previously proposed.
- Genuinely different regimes.
- Multiple transition variables.
- Long-range dependence and intermittent dynamics.
Main Contribution

Key Idea
- Use an unified tree-structured framework (model) to deal with structural breaks and regime-shifts.
- Combination of regression trees and smooth transition models.

Main Advantages
- Nests several nonlinear models previously proposed.
- Genuinely different regimes.
- Multiple transition variables.
- Long-range dependence and intermittent dynamics.
Main Contribution

Key Idea

- Use an unified tree-structured framework (model) to deal with structural breaks and regime-shifts.
- Combination of regression trees and smooth transition models.

Main Advantages

- Nests several nonlinear models previously proposed.
- Genuinely different regimes.
- Multiple transition variables.
- Long-range dependence and intermittent dynamics.
Main Results

Some questions

- Do volatility levels change in periods of significant losses or gains (cumulated returns)?
- Can negative returns over some horizon be associated with regimes of higher volatility?

Main findings

- New transition variable: Past cumulated returns.
- The effects of macroeconomic announcements and weekdays are also significant.
- Extremely good forecasts!
- Jumps do not improve the forecasts.
Some questions

- Do volatility levels change in periods of significant losses or gains (cumulated returns)?
- Can negative returns over some horizon be associated with regimes of higher volatility?

Main findings

- New transition variable: Past cumulated returns
- The effects of macroeconomic announcements and weekdays are also significant.
- Extremely good forecasts!
- Jumps do not improve the forecasts.
Some questions

- Do volatility levels change in periods of significant losses or gains (cumulated returns)?
- Can negative returns over some horizon be associated with regimes of higher volatility?

Main findings

- New transition variable: Past cumulated returns.
- The effects of macroeconomic announcements and weekdays are also significant.
- Extremely good forecasts!
- Jumps do not improve the forecasts.
Some questions

- Do volatility levels change in periods of significant losses or gains (cumulated returns)?
- Can negative returns over some horizon be associated with regimes of higher volatility?

Main findings

- **New transition variable:** Past cumulated returns.
- The effects of macroeconomic announcements and weekdays are also significant.
- Extremely good forecasts!
- Jumps do not improve the forecasts.
Main Results

Some questions

- Do volatility levels change in periods of significant losses or gains (cumulated returns)?
- Can negative returns over some horizon be associated with regimes of higher volatility?

Main findings

- New transition variable: Past cumulated returns.
- The effects of macroeconomic announcements and weekdays are also significant.
- Extremely good forecasts!
- Jumps do not improve the forecasts.
Some questions

- Do volatility levels change in periods of significant losses or gains (cumulated returns)?
- Can negative returns over some horizon be associated with regimes of higher volatility?

Main findings

- New transition variable: Past cumulated returns.
- The effects of macroeconomic announcements and weekdays are also significant.
- Extremely good forecasts!
- Jumps do not improve the forecasts.
Main Results

Some questions

- Do volatility levels change in periods of significant losses or gains (cumulated returns)?
- Can negative returns over some horizon be associated with regimes of higher volatility?

Main findings

- New transition variable: Past cumulated returns.
- The effects of macroeconomic announcements and weekdays are also significant.
- Extremely good forecasts!
- Jumps do not improve the forecasts.
Main Results

Example

IBM data: Monthly returns x Daily volatility

- Positive returns bring declines in the volatility
- High volatility regimes appear in more negative months
- NASDAQ Bubble Burst
- Log Realized Volatility
- Monthly Returns

Marcel Scharth and Marcelo C. Medeiros
Asymmetries and Long Memory in Volatility
Related literature

Persistence and regime switching

- Diebold and Inoue (Journal of Econometrics, 2001)
- Mikosch and Stărică (REStat, 2004)
- Hyung, Poon, and Granger (WP, 2005)
- Davidson and Sibbertsen (Journal of Econometrics, 2005)
- Hillebrand (Journal of Econometrics, 2005)

Nonlinear models and realized volatility

- Martens, van Dijk, and de Pooter (WP, 2004)
  - Long memory and nonlinearity
Related literature

Persistence and regime switching
- Diebold and Inoue (Journal of Econometrics, 2001)
- Mikosch and Stărică (REStat, 2004)
- Hyung, Poon, and Granger (WP, 2005)
- Davidson and Sibbertsen (Journal of Econometrics, 2005)
- Hillebrand (Journal of Econometrics, 2005)

Nonlinear models and realized volatility
- Martens, van Dijk, and de Pooter (WP, 2004)
  - Long memory and nonlinearity
Outline

1. The model
   - Overview and motivation
   - Mathematical definition
   - Modeling cycle

2. Empirical Results
   - Estimation results
   - Forecasting results

3. Conclusions
A parametric model based on the recursive partitioning of the covariate space $X$.

- A local model is determined in each of the $K \in \mathbb{N}$ different regions (partitions) of $X$.

- The model is displayed in a graph which has the format of a decision tree with $N \in \mathbb{N}$ parent (or split) nodes and $K \in \mathbb{N}$ terminal nodes (or leaves).

- Usually, the partitions are defined by a set of hyperplanes, each of which is orthogonal to the axis of a given predictor variable, called the split variable.

- Smooth transition between regimes.
Some Notation

- A parametric model based on the recursive partitioning of the covariate space $\mathbf{X}$.
  - A local model is determined in each of the $K \in \mathbb{N}$ different regions (partitions) of $\mathbf{X}$.

- The model is displayed in a graph which has the format of a decision tree with $N \in \mathbb{N}$ parent (or split) nodes and $K \in \mathbb{N}$ terminal nodes (or leaves).

- Usually, the partitions are defined by a set of hyperplanes, each of which is orthogonal to the axis of a given predictor variable, called the split variable.

- Smooth transition between regimes.
A parametric model based on the recursive partitioning of the covariate space $X$.

- A local model is determined in each of the $K \in \mathbb{N}$ different regions (partitions) of $X$.

The model is displayed in a graph which has the format of a decision tree with $N \in \mathbb{N}$ parent (or split) nodes and $K \in \mathbb{N}$ terminal nodes (or leaves).

- Usually, the partitions are defined by a set of hyperplanes, each of which is orthogonal to the axis of a given predictor variable, called the split variable.

- Smooth transition between regimes.
Some Notation

- A parametric model based on the recursive partitioning of the covariate space $\mathbb{X}$.
  - A local model is determined in each of the $K \in \mathbb{N}$ different regions (partitions) of $\mathbb{X}$.
- The model is displayed in a graph which has the format of a decision tree with $N \in \mathbb{N}$ parent (or split) nodes and $K \in \mathbb{N}$ terminal nodes (or leaves).
- Usually, the partitions are defined by a set of hyperplanes, each of which is orthogonal to the axis of a given predictor variable, called the split variable.
- Smooth transition between regimes.
Model Setup

Some Notation

- A parametric model based on the recursive partitioning of the covariate space $\mathbf{X}$.
  - A local model is determined in each of the $K \in \mathbb{N}$ different regions (partitions) of $\mathbf{X}$.

- The model is displayed in a graph which has the format of a decision tree with $N \in \mathbb{N}$ parent (or split) nodes and $K \in \mathbb{N}$ terminal nodes (or leaves).

- Usually, the partitions are defined by a set of hyperplanes, each of which is orthogonal to the axis of a given predictor variable, called the split variable.

- Smooth transition between regimes.
Model Setup

Some Notation (cont.)

- The root node is at position 0 and a parent node at position $j$ generates left- and right-child nodes at positions $2j + 1$ and $2j + 2$, respectively.
- Every parent node has an associated split variable $x_{sjt} \in x_t$, where $s_j \in S = \{1, 2, \ldots, q\}$.
- $\mathcal{J}$ and $\mathcal{T}$ are the sets of indexes of the parent and terminal nodes, respectively.
- A tree architecture can be fully determined by $\mathcal{J}$ and $\mathcal{T}$.
The root node is at position 0 and a parent node at position $j$ generates left- and right-child nodes at positions $2j + 1$ and $2j + 2$, respectively.

Every parent node has an associated split variable $x_{sjt} \in x_t$, where $s_j \in S = \{1, 2, \ldots, q\}$.

$J$ and $T$ are the sets of indexes of the parent and terminal nodes, respectively.

A tree architecture can be fully determined by $J$ and $T$. 
Some Notation (cont.)

- The root node is at position 0 and a parent node at position $j$ generates left- and right-child nodes at positions $2j + 1$ and $2j + 2$, respectively.
- Every parent node has an associated split variable $x_{s_jt} \in x_t$, where $s_j \in S = \{1, 2, \ldots, q\}$.
- $J$ and $T$ are the sets of indexes of the parent and terminal nodes, respectively.
- A tree architecture can be fully determined by $J$ and $T$. 
The root node is at position 0 and a parent node at position $j$ generates left- and right-child nodes at positions $2j + 1$ and $2j + 2$, respectively.

Every parent node has an associated split variable $x_{sjt} \in \mathbf{x}_t$, where $s_j \in \mathbb{S} = \{1, 2, \ldots, q\}$.

$\mathbb{J}$ and $\mathbb{T}$ are the sets of indexes of the parent and terminal nodes, respectively.

A tree architecture can be fully determined by $\mathbb{J}$ and $\mathbb{T}$. 
The root node is at position 0 and a parent node at position $j$ generates left- and right-child nodes at positions $2j + 1$ and $2j + 2$, respectively.

Every parent node has an associated split variable $x_{sjt} \in x_t$, where $s_j \in S = \{1, 2, \ldots, q\}$.

$\mathcal{J}$ and $\mathcal{T}$ are the sets of indexes of the parent and terminal nodes, respectively.

A tree architecture can be fully determined by $\mathcal{J}$ and $\mathcal{T}$. 
Model Setup

Example

\[
\begin{align*}
\sigma_t &= \omega_1 + \varepsilon_t \\
\sigma_t &= \omega_2 + \varepsilon_t \\
\sigma_t &= \omega_3 + \varepsilon_t
\end{align*}
\]
Model Setup

Mathematical Definition

- Let \( z_t \subseteq x_t \) such that \( z_t \in \mathbb{R}^p, p \leq q \). Set \( \tilde{z}_t = (1, z_t)' \). \( w_t \in \mathbb{R}^d \) is a vector of linear regressors.

- The Smooth Transition Regression Tree (STR-Tree) model (da Rosa, Veiga and Medeiros (WP, 2003)):

\[
\log(RV_t) = H_{JT}(x_t, w_t; \psi) + \varepsilon_t = \alpha' w_t + \sum_{i \in T} \beta_i' \tilde{z}_t B_{ji}(x_t; \theta_i) + \varepsilon_t
\]

where

\[
B_{ji}(x_t; \theta_i) = \prod_{j \in J} G(x_{sj,t}; \gamma_j, \nu_j)^{n_i,j(1+n_i,j)} \times [1 - G(x_{sj,t}; \gamma_j, \nu_j)]^{(1-n_i,j)(1+n_i,j)}
\]

and \( G(\cdot) \) is the logistic function.
Model Setup

Mathematical Definition

- Let $z_t \subseteq x_t$ such that $z_t \in \mathbb{R}^p$, $p \leq q$. Set $\tilde{z}_t = (1, z_t)'$. $w_t \in \mathbb{R}^d$ is a vector of linear regressors.

- The Smooth Transition Regression Tree (STR-Tree) model (da Rosa, Veiga and Medeiros (WP, 2003)):

$$
\log(RV_t) = H_{JT}(x_t, w_t; \psi) + \varepsilon_t = \alpha'w_t + \sum_{i \in T} \beta_i'\tilde{z}_tB_{ji}(x_t; \theta_i) + \varepsilon_t
$$

where

$$
B_{ji}(x_t; \theta_i) = \prod_{j \in J} G(x_{sj,t}; \gamma_j, c_j)^{\frac{n_{i,j}(1+n_{i,j})}{2}} \times
$$

$$
[1 - G(x_{sj,t}; \gamma_j, c_j)]^{(1-n_{i,j})(1+n_{i,j})}
$$

and $G(\cdot)$ is the logistic function.
Mathematical Definition

Let \( z_t \subseteq x_t \) such that \( z_t \in \mathbb{R}^p \), \( p \leq q \). Set \( \tilde{z}_t = (1, z_t)' \). \( w_t \in \mathbb{R}^d \) is a vector of linear regressors.

The Smooth Transition Regression Tree (STR-Tree) model (da Rosa, Veiga and Medeiros (WP, 2003)):

\[
\log(RV_t) = H_{JT}(x_t, w_t; \psi) + \varepsilon_t = \alpha' w_t + \sum_{i \in T} \beta^i_t \tilde{z}_t B_{ji}(x_t; \theta_i) + \varepsilon_t
\]

where

\[
B_{ji}(x_t; \theta_i) = \prod_{j \in J} G(x_{s_j}, t; \gamma_j, c_j)^{\frac{n_{i,j}(1+n_{i,j})}{2}} \times [1 - G(x_{s_j}, t; \gamma_j, c_j)]^{(1-n_{i,j})(1+n_{i,j})}
\]

and \( G(\cdot) \) is the logistic function.
The variable $n_{i,j}$ is defined as

$$n_{i,j} = \begin{cases} 
-1 & \text{if the path to leaf } i \text{ does not include the parent node } j; \\
0 & \text{if the path to leaf } i \text{ includes the right-child node of the parent node } j; \\
1 & \text{if the path to leaf } i \text{ includes the left-child node of the parent node } j. 
\end{cases}$$

Let $J_i$ be the subset of $J$ containing the indexes of the parent nodes that form the path to leaf $i$. Then, $\theta_i$ is the vector containing all the parameters $(\gamma_k, c_k)$ such that $k \in J_i, i \in T$. 

Marcel Scharth and Marcelo C. Medeiros

Asymmetries and Long Memory in Volatility
Mathematical Definition (cont.)

- The variable $n_{i,j}$ is defined as

$$n_{i,j} = \begin{cases} 
-1 & \text{if the path to leaf } i \text{ does not include the parent node } j; \\
0 & \text{if the path to leaf } i \text{ includes the right-child node} \\
of the parent node } j; \\
1 & \text{if the path to leaf } i \text{ includes the left-child node} \\
of the parent node } j. 
\end{cases}$$

- Let $\mathcal{J}_i$ be the subset of $\mathcal{J}$ containing the indexes of the parent nodes that form the path to leaf $i$. Then, $\theta_i$ is the vector containing all the parameters $(\gamma_k, c_k)$ such that $k \in \mathcal{J}_i$, $i \in \mathcal{T}$. 

Marcel Scharth and Marcelo C. Medeiros  
Asymmetries and Long Memory in Volatility
Mathematical Definition (cont.)

- The variable \( n_{i,j} \) is defined as

\[
  n_{i,j} = \begin{cases} 
  -1 & \text{if the path to leaf } i \text{ does not include the parent node } j; \\
  0 & \text{if the path to leaf } i \text{ includes the right-child node} \\
  1 & \text{of the parent node } j; \\
  & \text{of the parent node } j. 
  \end{cases}
\]

- Let \( \mathcal{J}_i \) be the subset of \( \mathcal{J} \) containing the indexes of the parent nodes that form the path to leaf \( i \). Then, \( \theta_i \) is the vector containing all the parameters \( (\gamma_k, c_k) \) such that \( k \in \mathcal{J}_i, i \in \mathbb{T} \).
STR-Tree Model Specification

Main Steps

- Following the “specific-to-general” principle, we start the cycle from the root node (depth 0). The general steps are:
  1. Selection of the relevant variables.
  2. Specification of the model by selecting in the depth $d$, using the LM test, a node to be split (if not in the root node) and a splitting variable.
  3. Estimation of the parameters.
  4. Evaluation of the estimated model by checking if it is necessary to:
     1. Change the node to be split.
     2. Change the splitting variable.
     3. Remove the split.
STR-Tree Model Specification

Main Steps

- Following the “specific-to-general” principle, we start the cycle from the root node (depth 0). The general steps are:
  1. **Selection of the relevant variables.**
  2. Specification of the model by selecting in the depth $d$, using the LM test, a node to be split (if not in the root node) and a splitting variable.
  3. Estimation of the parameters.
  4. Evaluation of the estimated model by checking if it is necessary to:
     1. Change the node to be split.
     2. Change the splitting variable.
     3. Remove the split.
STR-Tree Model Specification

Main Steps

- Following the “specific-to-general” principle, we start the cycle from the root node (depth 0). The general steps are:
  1. Selection of the relevant variables.
  2. Specification of the model by selecting in the depth $d$, using the LM test, a node to be split (if not in the root node) and a splitting variable.
  3. Estimation of the parameters.
  4. Evaluation of the estimated model by checking if it is necessary to:
     1. Change the node to be split.
     2. Change the splitting variable.
     3. Remove the split.
STR-Tree Model Specification

Main Steps

Following the “specific-to-general” principle, we start the cycle from the root node (depth 0). The general steps are:

1. Selection of the relevant variables.
2. Specification of the model by selecting in the depth \( d \), using the LM test, a node to be split (if not in the root node) and a splitting variable.
3. Estimation of the parameters.
4. Evaluation of the estimated model by checking if it is necessary to:
   1. Change the node to be split.
   2. Change the splitting variable.
   3. Remove the split.
Following the “specific-to-general” principle, we start the cycle from the root node (depth 0). The general steps are:

1. Selection of the relevant variables.
2. Specification of the model by selecting in the depth $d$, using the LM test, a node to be split (if not in the root node) and a splitting variable.
3. Estimation of the parameters.
4. Evaluation of the estimated model by checking if it is necessary to:
   1. Change the node to be split.
   2. Change the splitting variable.
   3. Remove the split.
Growing the Tree

Testing for an additional split

- Consider a STR-Tree model with $K$ leaves. We want to test if the terminal node $i^* \in \mathbb{T}$ should be split or not.

- Write the model as

$$
\log(RV_t) = \alpha' w_t + \sum_{i \in \mathbb{T} - \{i^*\}} \beta_i' \tilde{z}_t B_{ji} (x_t; \theta_i) +
\beta_{2i^*+1}' \tilde{z}_t B_{2i^*+1} (x_t; \theta_{2i^*+1}) + \beta_{2i^*+2}' \tilde{z}_t B_{2i^*+2} (x_t; \theta_{2i^*+2}) + \varepsilon_t,
$$

where

$$
B_{2i^*+1} (x_t; \theta_{2i^*+1}) = B_{i^*} (x_t; \theta_{i^*}) G(x_{i^*t}; \gamma_{i^*}, c_{i^*})
$$

$$
B_{2i^*+2} (x_t; \theta_{2i^*+2}) = B_{i^*} (x_t; \theta_{i^*}) [1 - G(x_{i^*t}; \gamma_{i^*}, c_{i^*})].
$$
Growing the Tree

Testing for an additional split

- Consider a STR-Tree model with $K$ leaves. We want to test if the terminal node $i^* \in \mathbb{T}$ should be split or not.

- Write the model as

$$
\log(RV_t) = \alpha' w_t + \sum_{i \in \mathbb{T} - \{i^*\}} \beta_i' z_t B_{ji}(x_t; \theta_i) +
\beta_{2i^*+1}' z_t B_{2i^*+1}(x_t; \theta_{2i^*+1}) + \beta_{2i^*+2}' z_t B_{2i^*+2}(x_t; \theta_{2i^*+2}) + \epsilon_t,
$$

where

$$
B_{2i^*+1}(x_t; \theta_{2i^*+1}) = B_{ji}(x_t; \theta_i) G(x_{i^*t}; \gamma_{i^*}, c_{i^*})
$$

$$
B_{2i^*+2}(x_t; \theta_{2i^*+2}) = B_{ji}(x_t; \theta_i) [1 - G(x_{i^*t}; \gamma_{i^*}, c_{i^*})].
$$
Testing for an additional split

Consider a STR-Tree model with $K$ leaves. We want to test if the terminal node $i^* \in \mathbb{T}$ should be split or not.

Write the model as

$$\log(RV_t) = \alpha' w_t + \sum_{i \in \mathbb{T} - \{i^*\}} \beta'_i \tilde{z}_t B_{ji} (x_t; \theta_i) + \beta'_{i^*+1} \tilde{z}_t B_{2i^*+1} (x_t; \theta_{2i^*+1}) + \beta'_{i^*+2} \tilde{z}_t B_{2i^*+2} (x_t; \theta_{2i^*+2}) + \epsilon_t,$$

where

$$B_{2i^*+1} (x_t; \theta_{2i^*+1}) = B_{ji} (x_t; \theta_i^*) G(x_i^* t; \gamma_i^*, c_i^*)$$

$$B_{2i^*+2} (x_t; \theta_{2i^*+2}) = B_{ji} (x_t; \theta_i^*) [1 - G(x_i^* t; \gamma_i^*, c_i^*)].$$
Growing the tree

Testing for an additional split (cont.)

- In a more compact form

\[
\log(RV_t) = \alpha' w_t + \sum_{i \in T - \{i^*\}} \beta'_i \tilde{z}_t B_{ji} (x_t; \theta_i) + \phi' \tilde{z}_t B_{j^*i^*} (x_t; \theta_i^*) + \lambda' \tilde{z}_t B_{j^*i^*} (x_t; \theta_i^*) G(x_{i^*t}; \gamma_{i^*}, c_{i^*}) + \varepsilon_t,
\]

where \( \phi = \beta_{2i^*+2} \) and \( \lambda = \beta_{2i^*+1} - \beta_{2i^*+2} \).

- In order to test the statistical significance of the split, a convenient null hypothesis is \( H_0 : \gamma_{i^*} = 0 \) against the alternative \( H_a : \gamma_{i^*} > 0 \).
Growing the tree

Testing for an additional split (cont.)

- In a more compact form

$$\log(RV_t) = \alpha' w_t + \sum_{i \in T - \{i^*\}} \beta_i' \tilde{z}_t B_{ji} (x_t; \theta_i) + \phi' \tilde{z}_t B_{ji^*} (x_t; \theta_{i^*}) + \lambda' \tilde{z}_t B_{ji^*} (x_t; \theta_{i^*}) G(x_{i^* t}; \gamma_{i^*}, c_{i^*}) + \varepsilon_t,$$

where $\phi = \beta_{2i^*+2}$ and $\lambda = \beta_{2i^*+1} - \beta_{2i^*+2}$.

- In order to test the statistical significance of the split, a convenient null hypothesis is $H_0 : \gamma_{i^*} = 0$ against the alternative $H_a : \gamma_{i^*} > 0$. 
Growing the tree

Testing for an additional split (cont.)

- In a more compact form

\[
\log(RV_t) = \alpha' w_t + \sum_{i \in T - \{i^*\}} \beta_i' \tilde{z}_t B_{j_i} (x_t; \theta_i) + \\
\phi' \tilde{z}_t B_{j_i^*} (x_t; \theta_{i^*}) + \lambda' \tilde{z}_t B_{j_i^*} (x_t; \theta_{i^*}) \ G(x_{i^*t}; \gamma_{i^*}, c_{i^*}) + \varepsilon_t,
\]

where \( \phi = \beta_{2i^*+2} \) and \( \lambda = \beta_{2i^*+1} - \beta_{2i^*+2} \).

- In order to test the statistical significance of the split, a convenient null hypothesis is \( H_0 : \gamma_{i^*} = 0 \) against the alternative \( H_a : \gamma_{i^*} > 0 \).
Identification Problem

Under $H_0$, the parameters $\lambda$ and $c_{t^*}$ can assume different values without changing the quasi-loglikelihood function.

Solution

- Third-order Taylor expansion around $\gamma_{i^*} = 0$.

$$
\log(RV_t) = \alpha' w_t + \sum_{i \in T - \{i^*\}} \beta_i' \bar{Z}_t B_{ji} (x_t; \theta_i) + \alpha_0' \bar{Z}_t B_{ji^*} (x_t; \theta_{i^*}) + \alpha_1' \bar{Z}_t B_{ji^*} (x_t; \theta_{i^*}) x_{i^* t} + \alpha_2' \bar{Z}_t B_{ji^*} (x_t; \theta_{i^*}) x_{i^* t}^2 + \alpha_3' \bar{Z}_t B_{ji^*} (x_t; \theta_{i^*}) x_{i^* t}^3 + e_t,
$$

where $e_t = \varepsilon_t + \lambda' \bar{Z}_t B_{ji^*} (x_t; \theta_{i^*}) \times \text{Remainder}$.

- Thus the null hypothesis becomes $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$. 
Growing the Tree

Identification Problem

Under $H_0$, the parameters $\lambda$ and $c_{i*}$ can assume different values without changing the quasi-loglikelihood function.

Solution

- **Third-order Taylor expansion around $\gamma_{i*} = 0$.**

\[
\log(RV_t) = \alpha' w_t + \sum_{i \in T \setminus \{i^*\}} \beta_i' \tilde{z}_t B_{i*} (x_t; \theta_i) + \alpha_0' \tilde{z}_t B_{i*} (x_t; \theta_{i*}) + \\
\alpha_1' \tilde{z}_t B_{i*} (x_t; \theta_{i*}) x_{i* t} + \alpha_2' \tilde{z}_t B_{i*} (x_t; \theta_{i*}) x_{i* t}^2 + \\
\alpha_3' \tilde{z}_t B_{i*} (x_t; \theta_{i*}) x_{i* t}^3 + e_t,
\]

where $e_t = \varepsilon_t + \lambda' \tilde{z}_t B_{i*} (x_t; \theta_{i*}) \times \text{Remainder}.$

- Thus the null hypothesis becomes $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$. 

Growing the Tree

Identification Problem

Under $H_0$, the parameters $\lambda$ and $c_{i*}$ can assume different values without changing the quasi-loglikelihood function.

Solution

Third-order Taylor expansion around $\gamma_{i*} = 0$.

$$
\log(RV_t) = \alpha' w_t + \sum_{i \in T - \{i^*\}} \beta'_i \tilde{z}_t B_{ij} (x_t; \theta_i) + \alpha'_0 \tilde{z}_t B_{ij*} (x_t; \theta_{i*}) + \alpha'_1 \tilde{z}_t B_{ij*} (x_t; \theta_{i*}) x_{i* t} + \alpha'_2 \tilde{z}_t B_{ij*} (x_t; \theta_{i*}) x_{i* t}^2 + \alpha'_3 \tilde{z}_t B_{ij*} (x_t; \theta_{i*}) x_{i* t}^3 + \epsilon_t,
$$

where $\epsilon_t = \epsilon_t + \lambda' \tilde{z}_t B_{ij*} (x_t; \theta_{i*}) \times \text{Remainder}$.

Thus the null hypothesis becomes $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$. 
A robust version of the test can be carried out in stages as follows:

1. Estimate the STR-Tree model with $K$ regimes and save the residuals $\hat{\varepsilon}_t$.
2. Regress $\hat{\nu}_t$ on $\hat{h}_t$ and estimate the residuals $\hat{r}_t$, where
   \[
   \hat{h}_t = \left. \frac{\partial H_{JT}(x_t; \psi)}{\partial \psi} \right|_{\psi = \hat{\psi}}
   \quad \text{and} \quad
   \hat{\nu}_t = \left[ \hat{z}_t \hat{B}_{ji} x_{i^* t}, \hat{z}_t \hat{B}_{ji} x_{i^* t}^2, \hat{z}_t \hat{B}_{ji} x_{i^* t}^3 \right]
   \]
3. Regress a vector of ones on $\hat{\varepsilon}_t\hat{r}_t$ and compute the sum of the squared residuals $SSR$. Compute the $LM$ statistic
   \[
   LM = T - SSR \xrightarrow{d} \chi^2(\text{dim}(\nu_t))
   \]
Growing the tree

Testing for an additional split (cont.)

A robust version of the test can be carried out in stages as follows:

1. **Estimate the STR-Tree model with $K$ regimes and save the residuals $\hat{\varepsilon}_t$.**

2. **Regress $\hat{\nu}_t$ on $\hat{h}_t$ and estimate the residuals $\hat{r}_t$, where**

   $$\hat{h}_t = \left. \frac{\partial H_{JT}(x_t; \psi)}{\partial \psi} \right|_{\psi = \hat{\psi}}$$

   and

   $$\hat{\nu}_t = \left[ \hat{z}_t \hat{B}_{ij} x_{i_s}^t, \hat{z}_t \hat{B}_{ij} x_{i_s}^2 t, \hat{z}_t \hat{B}_{ij} x_{i_s}^3 t \right]$$

3. **Regress a vector of ones on $\hat{\varepsilon}_t \hat{r}_t$ and compute the sum of the squared residuals $SSR$. Compute the $LM$ statistic**

   $$LM = T - SSR \xrightarrow{d} \chi^2(\dim(\nu_t))$$
Growing the tree

Testing for an additional split (cont.)

A robust version of the test can be carried out in stages as follows:

1. Estimate the STR-Tree model with $K$ regimes and save the residuals $\hat{\varepsilon}_t$.
2. Regress $\hat{\nu}_t$ on $\hat{h}_t$ and estimate the residuals $\hat{r}_t$, where

$$\hat{h}_t = \left. \frac{\partial H_{JT}(x_t; \psi)}{\partial \psi} \right|_{\psi=\hat{\psi}}$$

and

$$\hat{\nu}_t = \left[ \bar{z}_t \hat{B}_i x_{i*}^1, \bar{z}_t \hat{B}_i x_{i*}^2, \bar{z}_t \hat{B}_i x_{i*}^3 \right]$$

3. Regress a vector of ones on $\hat{\varepsilon}_t \hat{r}_t$ and compute the sum of the squared residuals SSR. Compute the $LM$ statistic

$$LM = T - SSR \overset{d}{\to} \chi^2(\text{dim}(\nu_t))$$
Growing the tree

Testing for an additional split (cont.)

A robust version of the test can be carried out in stages as follows:

1. Estimate the STR-Tree model with $K$ regimes and save the residuals $\hat{\varepsilon}_t$.
2. Regress $\hat{\nu}_t$ on $\hat{h}_t$ and estimate the residuals $\hat{r}_t$, where

$$
\hat{h}_t = \left. \frac{\partial H_{JT}(x_t; \psi)}{\partial \psi} \right|_{\psi = \hat{\psi}}
$$

and

$$
\hat{\nu}_t = \begin{bmatrix}
\tilde{z}_t \hat{B}_{i^*} x_{i^* t}^1 \\
\tilde{z}_t \hat{B}_{i^*} x_{i^* t}^2 \\
\tilde{z}_t \hat{B}_{i^*} x_{i^* t}^3
\end{bmatrix}
$$

3. Regress a vector of ones on $\hat{\varepsilon}_t \hat{r}_t$ and compute the sum of the squared residuals $SSR$. Compute the $LM$ statistic

$$
LM = T - SSR \xrightarrow{d} \chi^2(\text{dim}(\nu_t))
$$
Data

Description

- 16 DJIA stocks: Alcoa (AA), American International Group (AIG), Boeing (BA), Caterpillar (CAT), General Electric (GE), General Motors (GM), Hewlett Packard (HPQ), IBM (IBM), Intel (INTC), Johnson and Johnson (JNJ), Coca-Cola (KO), Microsoft (MSFT), Merck (MRK), Pfizer (PFE), Wal-Mart (WMT) and Exxon (XON).


- Days with abnormally small trading volume are excluded.
Data

Description

- 16 DJIA stocks: Alcoa (AA), American International Group (AIG), Boeing (BA), Caterpillar (CAT), General Electric (GE), General Motors (GM), Hewlett Packard (HPQ), IBM (IBM), Intel (INTC), Johnson and Johnson (JNJ), Coca-Cola (KO), Microsoft (MSFT), Merck (MRK), Pfizer (PFE), Wal-Mart (WMT) and Exxon (XON).


- Days with abnormally small trading volume are excluded.
Data

Description

- 16 DJIA stocks: Alcoa (AA), American International Group (AIG), Boeing (BA), Caterpillar (CAT), General Electric (GE), General Motors (GM), Hewlett Packard (HPQ), IBM (IBM), Intel (INTC), Johnson and Johnson (JNJ), Coca-Cola (KO), Microsoft (MSFT), Merck (MRK), Pfizer (PFE), Wal-Mart (WMT) and Exxon (XON).


- Days with abnormally small trading volume are excluded.
Data

Description

- 16 DJIA stocks: Alcoa (AA), American International Group (AIG), Boeing (BA), Caterpillar (CAT), General Electric (GE), General Motors (GM), Hewlett Packard (HPQ), IBM (IBM), Intel (INTC), Johnson and Johnson (JNJ), Coca-Cola (KO), Microsoft (MSFT), Merck (MRK), Pfizer (PFE), Wal-Mart (WMT) and Exxon (XON).


- Days with abnormally small trading volume are excluded.
Pre-processing and volatility estimation

- Non-standard quotes removal and computation of mid-quote prices ⇒ one second returns.
- Following Hansen and Lunde (2006), the previous tick method is used for determining prices at precise time marks.
- Realized volatility is constructed with the two time scales estimator with five-minute grids.
Pre-processing and volatility estimation

- Non-standard quotes removal and computation of mid-quote prices ⇒ one second returns.
- Following Hansen and Lunde (2006), the previous tick method is used for determining prices at precise time marks.
- Realized volatility is constructed with the two time scales estimator with five-minute grids.
Pre-processing and volatility estimation

- Non-standard quotes removal and computation of mid-quote prices ⇒ one second returns.
- Following Hansen and Lunde (2006), the previous tick method is used for determining prices at precise time marks.
- Realized volatility is constructed with the two time scales estimator with five-minute grids.
Pre-processing and volatility estimation

- Non-standard quotes removal and computation of mid-quote prices ⇒ one second returns.
- Following Hansen and Lunde (2006), the previous tick method is used for determining prices at precise time marks.
- Realized volatility is constructed with the two time scales estimator with five-minute grids.
### The STR-Tree model

The estimated STR-Tree model has the following structure:

$$
\log(RV_t) = \sum_{i \in \mathbb{T}} \beta_i B_{ji}(x_t; \theta_i) + \alpha_1 \log(RV_{t-1}) + \cdots + \alpha_p \log(RV_{t-p}) + \\
\delta_1 I[Mon]_t + \delta_2 I[Tue]_t + \delta_3 I[Wed]_t + \delta_4 I[Thu]_t + \\
\delta_5 I[FOMC]_t + \delta_6 I[EMP]_t + \delta_7 I[CPI]_t + \delta_8 I[PPI]_t + \varepsilon_t
$$

- $I[Mon]_t$, $I[Tue]_t$, $I[Wed]_t$, and $I[Thu]_t$ are dummy variables for the weekdays.
- $I[FOMC]_t$, $I[EMP]_t$, $I[CPI]_t$, and $I[PPI]_t$ are dummy variables for the announcement dates.
- $x_t$ contains lagged cumulated returns over the one to 120 days.
The estimated STR-Tree model has the following structure:

\[
\log(RV_t) = \sum_{i \in T} \beta_i B_{ij}(x_t; \theta_i) + \alpha_1 \log(RV_{t-1}) + \cdots + \alpha_p \log(RV_{t-p}) + \\
\delta_1 I[Mon]_t + \delta_2 I[Tue]_t + \delta_3 I[Wed]_t + \delta_4 I[Thu]_t + \\
\delta_5 I[FOMC]_t + \delta_6 I[EMP]_t + \delta_7 I[CPI]_t + \delta_8 I[PPI]_t + \varepsilon_t
\]

- \(I[Mon]_t, I[Tue]_t, I[Wed]_t, \text{ and } I[Thu]_t\) are dummy variables for the weekdays.
- \(I[FOMC]_t, I[EMP]_t, I[CPI]_t, \text{ and } I[PPI]_t\) are dummy variables for the announcement dates.
- \(x_t\) contains lagged cumulated returns over the one to 120 days.
The estimated STR-Tree model has the following structure:

\[
\log(RV_t) = \sum_{i \in T} \beta_i B_{ij}(x_t; \theta_i) + \alpha_1 \log(RV_{t-1}) + \cdots + \alpha_p \log(RV_{t-p}) + \\
\delta_1 I[Mon]_t + \delta_2 I[Tue]_t + \delta_3 I[Wed]_t + \delta_4 I[Thu]_t + \\
\delta_5 I[FOMC]_t + \delta_6 I[EMP]_t + \delta_7 I[CPI]_t + \delta_8 I[PPI]_t + \epsilon_t
\]

- \( I[Mon]_t, I[Tue]_t, I[Wed]_t, \) and \( I[Thu]_t \) are dummy variables for the weekdays.
- \( I[FOMC]_t, I[EMP]_t, I[CPI]_t, \) and \( I[PPI]_t \) are dummy variables for the announcement dates.
- \( x_t \) contains lagged cumulated returns over the one to 120 days.
The STR-Tree model

The estimated STR-Tree model has the following structure

\[
\log(RV_t) = \sum_{i \in T} \beta_i B_{ji}(x_t; \theta_i) + \alpha_1 \log(RV_{t-1}) + \cdots + \alpha_p \log(RV_{t-p}) + \\
\delta_1 I[Mon]_t + \delta_2 I[Tue]_t + \delta_3 I[Wed]_t + \delta_4 I[Thu]_t + \\
\delta_5 I[FOMC]_t + \delta_6 I[EMP]_t + \delta_7 I[CPI]_t + \delta_8 I[PPI]_t + \varepsilon_t
\]

- \(I[Mon]_t, I[Tue]_t, I[Wed]_t, \text{ and } I[Thu]_t\) are dummy variables for the weekdays.
- \(I[FOMC]_t, I[EMP]_t, I[CPI]_t, \text{ and } I[PPI]_t\) are dummy variables for the announcement dates.
- \(x_t\) contains lagged cumulated returns over the one to 120 days.
The estimated STR-Tree model has the following structure

\[ \log(RV_t) = \sum_{i \in T} \beta_i B_{i}(x_t; \theta_i) + \alpha_1 \log(RV_{t-1}) + \cdots + \alpha_p \log(RV_{t-p}) + \]

\[ + \delta_1 I[Mon]_t + \delta_2 I[Tue]_t + \delta_3 I[Wed]_t + \delta_4 I[Thu]_t + \]

\[ + \delta_5 I[FOMC]_t + \delta_6 I[EMP]_t + \delta_7 I[CPI]_t + \delta_8 I[PPI]_t + \varepsilon_t \]

- \( I[Mon]_t, I[Tue]_t, I[Wed]_t, \) and \( I[Thu]_t \) are dummy variables for the weekdays.
- \( I[FOMC]_t, I[EMP]_t, I[CPI]_t, \) and \( I[PPI]_t \) are dummy variables for the announcement dates.
- \( x_t \) contains lagged cumulated returns over the one to 120 days.
Estimated Models

Example

The estimated STR-Tree model: The case of IBM

\[
\begin{align*}
&\text{0} \\
&\text{1} \quad \begin{cases} 
  r_{90,t-1} \geq 9.3 \\
  r_{90,t-1} < 9.3
\end{cases} \\
&\text{2} \quad \begin{cases} 
  r_{39,t-1} \geq -11.9 \\
  r_{39,t-1} < -11.9
\end{cases} \\
&\text{5} \quad \begin{cases} 
  r_{5,t-1} \geq 2.26 \\
  r_{5,t-1} < 2.26
\end{cases} \\
&\text{6} \\
&\text{11} \quad \begin{cases} 
  r_{2,t-1} \geq -3.34 \\
  r_{2,t-1} < -3.34
\end{cases} \\
&\text{12} \\
&\text{23} \\
&\text{24}
\end{align*}
\]
Estimated Models

Alternative models

Apart from the STR-Tree model, the following models are also estimated:

- A structural break model (a STR-Tree specification with time as the only transition variable).
- Linear AR and ARFIMA models.
- The Heterogeneous Autoregressive (HAR) model put forward by Corsi (2003)
- The GARCH(1,1) model
- The Exponentially Weighted Moving Average (EWMA)
Estimated Models

Alternative models

- Apart from the STR-Tree model, the following models are also estimated:
  - A structural break model (a STR-Tree specification with time as the only transition variable).
  - Linear AR and ARFIMA models.
  - The Heterogeneous Autoregressive (HAR) model put forward by Corsi (2003)
  - The GARCH(1,1) model
  - The Exponentially Weighted Moving Average (EWMA)
Apart from the STR-Tree model, the following models are also estimated:

- A structural break model (a STR-Tree specification with time as the only transition variable).
- Linear AR and ARFIMA models.
- The Heterogeneous Autoregressive (HAR) model put forward by Corsi (2003)
- The GARCH(1,1) model
- The Exponentially Weighted Moving Average (EWMA)
Estimated Models

Alternative models

- Apart from the STR-Tree model, the following models are also estimated:
  - A structural break model (a STR-Tree specification with time as the only transition variable).
  - Linear AR and ARFIMA models.
  - The Heterogeneous Autoregressive (HAR) model put forward by Corsi (2003)
  - The GARCH(1,1) model
  - The Exponentially Weighted Moving Average (EWMA)
Estimated Models

Alternative models

- Apart from the STR-Tree model, the following models are also estimated:
  - A structural break model (a STR-Tree specification with time as the only transition variable).
  - Linear AR and ARFIMA models.
  - The Heterogeneous Autoregressive (HAR) model put forward by Corsi (2003)
    - The GARCH(1,1) model
    - The Exponentially Weighted Moving Average (EWMA)
      - Note that the EWMA is computed on the log realized volatility series with decay parameter 0.8.
Apart from the STR-Tree model, the following models are also estimated:

- A structural break model (a STR-Tree specification with time as the only transition variable).
- Linear AR and ARFIMA models.
- The Heterogeneous Autoregressive (HAR) model put forward by Corsi (2003)
- The GARCH(1,1) model
- The Exponentially Weighted Moving Average (EWMA)

Note that the EWMA is computed on the log realized volatility series with decay parameter 0.8.
Apart from the STR-Tree model, the following models are also estimated:

- A structural break model (a STR-Tree specification with time as the only transition variable).
- Linear AR and ARFIMA models.
- The Heterogeneous Autoregressive (HAR) model put forward by Corsi (2003)
- The GARCH(1,1) model
- The Exponentially Weighted Moving Average (EWMA)

Note that the EWMA is computed on the log realized volatility series with decay parameter 0.8.
Estimated Models

Alternative models

- Apart from the STR-Tree model, the following models are also estimated:
  - A structural break model (a STR-Tree specification with time as the only transition variable).
  - Linear AR and ARFIMA models.
  - The Heterogeneous Autoregressive (HAR) model put forward by Corsi (2003)
  - The GARCH(1,1) model
  - The Exponentially Weighted Moving Average (EWMA)
    - Note that the EWMA is computed on the log realized volatility series with decay parameter 0.8.
Main Setup

- Out-of-sample period: 01-Jan-2000 to 31-Dec-2003 (983 observations)
- Each model is re-estimated daily and then used for point and value at risk forecasting for the horizons of one, five, ten and twenty days ahead.
- The specification of the STR-Tree model is revised monthly.
- Point forecasts for the STR-Tree model are calculated through conditional simulation, as well as interval forecasts for all models.
Main Setup

- **Out-of-sample period**: 01-Jan-2000 to 31-Dec-2003 (983 observations)

- Each model is re-estimated daily and then used for point and value at risk forecasting for the horizons of one, five, ten and twenty days ahead.

- The specification of the STR-Tree model is revised monthly.

- Point forecasts for the STR-Tree model are calculated through conditional simulation, as well as interval forecasts for all models.
Main Setup

- Out-of-sample period: 01-Jan-2000 to 31-Dec-2003 (983 observations)
- Each model is re-estimated daily and then used for point and value at risk forecasting for the horizons of one, five, ten and twenty days ahead.
- The specification of the STR-Tree model is revised monthly.
- Point forecasts for the STR-Tree model are calculated through conditional simulation, as well as interval forecasts for all models.
Main Setup

- Out-of-sample period: 01-Jan-2000 to 31-Dec-2003 (983 observations)
- Each model is re-estimated daily and then used for point and value at risk forecasting for the horizons of one, five, ten and twenty days ahead.
- The specification of the STR-Tree model is revised monthly.
- Point forecasts for the STR-Tree model are calculated through conditional simulation, as well as interval forecasts for all models.
Main Setup

- Out-of-sample period: 01-Jan-2000 to 31-Dec-2003 (983 observations)
- Each model is re-estimated daily and then used for point and value at risk forecasting for the horizons of one, five, ten and twenty days ahead.
- The specification of the STR-Tree model is revised monthly.
- Point forecasts for the STR-Tree model are calculated through conditional simulation, as well as interval forecasts for all models.
### MAE and Forecasting Accuracy Tests

<table>
<thead>
<tr>
<th></th>
<th>1 day</th>
<th></th>
<th></th>
<th>5 days</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAE</td>
<td>HLN</td>
<td>SPA</td>
<td>MAE</td>
<td>HLN</td>
<td>SPA</td>
</tr>
<tr>
<td>STR-Tree/AE</td>
<td>0.322</td>
<td>0.000</td>
<td>0.960</td>
<td>0.397</td>
<td>0.000</td>
<td>0.975</td>
</tr>
<tr>
<td>STR-Tree/SB</td>
<td>0.365</td>
<td>0.000</td>
<td>0.000</td>
<td>0.474</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>STR-Tree/DJIA</td>
<td>0.324</td>
<td>0.000</td>
<td>0.456</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>STR-Tree/SB+AE</td>
<td>0.340</td>
<td>0.485</td>
<td>0.004</td>
<td>0.409</td>
<td>0.185</td>
<td>0.285</td>
</tr>
<tr>
<td>HAR</td>
<td>0.332</td>
<td>0.027</td>
<td>0.026</td>
<td>0.412</td>
<td>0.338</td>
<td>0.038</td>
</tr>
<tr>
<td>ARFIMA</td>
<td>0.339</td>
<td>–</td>
<td>0.001</td>
<td>0.414</td>
<td>–</td>
<td>0.032</td>
</tr>
<tr>
<td>AR</td>
<td>0.334</td>
<td>0.092</td>
<td>0.001</td>
<td>0.410</td>
<td>0.215</td>
<td>0.020</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.348</td>
<td>0.031</td>
<td>0.001</td>
<td>0.407</td>
<td>0.098</td>
<td>0.517</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.490</td>
<td>0.000</td>
<td>0.000</td>
<td>0.527</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>MAE</td>
<td>HLN</td>
<td>SPA</td>
<td>MAE</td>
<td>HLN</td>
<td>SPA</td>
</tr>
<tr>
<td>STR-Tree/AE</td>
<td>0.447</td>
<td>0.003</td>
<td>0.969</td>
<td>0.507</td>
<td>0.012</td>
<td>0.982</td>
</tr>
<tr>
<td>STR-Tree/SB</td>
<td>0.532</td>
<td>0.000</td>
<td>0.000</td>
<td>0.604</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>STR-Tree/DJIA</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>STR-Tree/SB+AE</td>
<td>0.460</td>
<td>0.321</td>
<td>0.446</td>
<td>0.510</td>
<td>0.025</td>
<td>0.890</td>
</tr>
<tr>
<td>HAR</td>
<td>0.466</td>
<td>0.311</td>
<td>0.039</td>
<td>0.535</td>
<td>0.160</td>
<td>0.001</td>
</tr>
<tr>
<td>ARFIMA</td>
<td>0.463</td>
<td>–</td>
<td>0.287</td>
<td>0.524</td>
<td>–</td>
<td>0.489</td>
</tr>
<tr>
<td>AR</td>
<td>0.458</td>
<td>0.249</td>
<td>0.131</td>
<td>0.518</td>
<td>0.253</td>
<td>0.269</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.463</td>
<td>0.473</td>
<td>0.390</td>
<td>0.536</td>
<td>0.090</td>
<td>0.233</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.555</td>
<td>0.000</td>
<td>0.000</td>
<td>0.591</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### Forecasting Results: IBM

#### Results by year: MAE for the years 2000–2002

<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Day</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA</td>
<td>0.459</td>
<td>0.373</td>
<td>0.352</td>
</tr>
<tr>
<td>STR-Tree</td>
<td>0.451</td>
<td>0.350</td>
<td>0.328</td>
</tr>
<tr>
<td>5 Days</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA</td>
<td>0.536</td>
<td>0.465</td>
<td>0.454</td>
</tr>
<tr>
<td>STR-Tree</td>
<td>0.547</td>
<td>0.420</td>
<td>0.428</td>
</tr>
<tr>
<td>10 Days</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA</td>
<td>0.567</td>
<td>0.537</td>
<td>0.525</td>
</tr>
<tr>
<td>STR-Tree</td>
<td>0.608</td>
<td>0.485</td>
<td>0.479</td>
</tr>
<tr>
<td>20 Days</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA</td>
<td>0.605</td>
<td>0.634</td>
<td>0.583</td>
</tr>
<tr>
<td>STR-Tree</td>
<td>0.633</td>
<td>0.567</td>
<td>0.529</td>
</tr>
</tbody>
</table>
## Forecasting Results: IBM

### Results for 2003 – One and 20 Days Ahead

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE 1 day</th>
<th>HLN 1 day</th>
<th>SPA 1 day</th>
<th>MAE 20 days</th>
<th>HLN 20 days</th>
<th>SPA 20 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>STR-Tree/AE</td>
<td>0.157</td>
<td>0.002</td>
<td>0.907</td>
<td>0.236</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>STR-Tree/SB</td>
<td>0.201</td>
<td>0.000</td>
<td>0.000</td>
<td>0.524</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>STR-Tree/DJIA</td>
<td>0.161</td>
<td>0.028</td>
<td>0.001</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>STR-Tree/SB+AE</td>
<td>0.165</td>
<td>0.169</td>
<td>0.010</td>
<td>0.297</td>
<td>0.005</td>
<td>0.000</td>
</tr>
<tr>
<td>HAR</td>
<td>0.156</td>
<td>0.000</td>
<td>0.951</td>
<td>0.191</td>
<td>0.000</td>
<td>0.848</td>
</tr>
<tr>
<td>ARFIMA</td>
<td>0.170</td>
<td>–</td>
<td>0.000</td>
<td>0.274</td>
<td>–</td>
<td>0.000</td>
</tr>
<tr>
<td>AR</td>
<td>0.159</td>
<td>0.000</td>
<td>0.011</td>
<td>0.207</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.158</td>
<td>0.005</td>
<td>0.695</td>
<td>0.200</td>
<td>0.000</td>
<td>0.539</td>
</tr>
<tr>
<td>GARCH</td>
<td>0.322</td>
<td>0.000</td>
<td>0.000</td>
<td>0.527</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
## Forecasting Results: All Series – One Day Ahead

<table>
<thead>
<tr>
<th>Series</th>
<th>STR-Tree/AE</th>
<th>STR-Tree/SB</th>
<th>STR-Tree/AE+SB</th>
<th>ARFIMA</th>
<th>HAR</th>
<th>EWMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>0.450</td>
<td>0.476</td>
<td>0.560</td>
<td>0.456</td>
<td>0.474</td>
<td>0.465</td>
</tr>
<tr>
<td></td>
<td>(0.933)</td>
<td>(0.001)</td>
<td>(0.067)</td>
<td>(0.301)</td>
<td>(0.000)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>AIG</td>
<td>0.359</td>
<td>0.371</td>
<td>0.372</td>
<td>0.364</td>
<td>0.369</td>
<td>0.371</td>
</tr>
<tr>
<td></td>
<td>(0.913)</td>
<td>(0.003)</td>
<td>(0.005)</td>
<td>(0.260)</td>
<td>(0.041)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>BA</td>
<td>0.393</td>
<td>0.414</td>
<td>0.404</td>
<td>0.397</td>
<td>0.405</td>
<td>0.409</td>
</tr>
<tr>
<td></td>
<td>(0.837)</td>
<td>(0.002)</td>
<td>(0.057)</td>
<td>(0.469)</td>
<td>(0.099)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>CAT</td>
<td>0.398</td>
<td>0.423</td>
<td>0.423</td>
<td>0.404</td>
<td>0.412</td>
<td>0.411</td>
</tr>
<tr>
<td></td>
<td>(0.904)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.152)</td>
<td>(0.044)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>GE</td>
<td>0.340</td>
<td>0.369</td>
<td>0.363</td>
<td>0.349</td>
<td>0.355</td>
<td>0.361</td>
</tr>
<tr>
<td></td>
<td>(0.873)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.118)</td>
<td>(0.008)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>GM</td>
<td>0.374</td>
<td>0.409</td>
<td>0.388</td>
<td>0.380</td>
<td>0.388</td>
<td>0.389</td>
</tr>
<tr>
<td></td>
<td>(0.920)</td>
<td>(0.000)</td>
<td>(0.003)</td>
<td>(0.181)</td>
<td>(0.003)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>HP</td>
<td>0.574</td>
<td>0.604</td>
<td>0.585</td>
<td>0.579</td>
<td>0.599</td>
<td>0.583</td>
</tr>
<tr>
<td></td>
<td>(0.903)</td>
<td>(0.000)</td>
<td>(0.084)</td>
<td>(0.372)</td>
<td>(0.002)</td>
<td>(0.314)</td>
</tr>
<tr>
<td>INTC</td>
<td>0.436</td>
<td>0.490</td>
<td>0.443</td>
<td>0.448</td>
<td>0.459</td>
<td>0.466</td>
</tr>
<tr>
<td></td>
<td>(0.821)</td>
<td>(0.000)</td>
<td>(0.331)</td>
<td>(0.055)</td>
<td>(0.004)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.368</td>
<td>0.380</td>
<td>0.385</td>
<td>0.372</td>
<td>0.379</td>
<td>0.380</td>
</tr>
<tr>
<td></td>
<td>(0.806)</td>
<td>(0.015)</td>
<td>(0.008)</td>
<td>(0.579)</td>
<td>(0.139)</td>
<td>(0.113)</td>
</tr>
<tr>
<td>KO</td>
<td>0.335</td>
<td>0.360</td>
<td>0.346</td>
<td>0.341</td>
<td>0.348</td>
<td>0.339</td>
</tr>
<tr>
<td></td>
<td>(0.904)</td>
<td>(0.000)</td>
<td>(0.014)</td>
<td>(0.164)</td>
<td>(0.006)</td>
<td>(0.473)</td>
</tr>
<tr>
<td>MRK</td>
<td>0.367</td>
<td>0.389</td>
<td>0.377</td>
<td>0.370</td>
<td>0.381</td>
<td>0.378</td>
</tr>
<tr>
<td></td>
<td>(0.886)</td>
<td>(0.001)</td>
<td>(0.034)</td>
<td>(0.634)</td>
<td>(0.010)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.347</td>
<td>0.380</td>
<td>0.364</td>
<td>0.357</td>
<td>0.363</td>
<td>0.369</td>
</tr>
<tr>
<td></td>
<td>(0.827)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.133)</td>
<td>(0.013)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>PFE</td>
<td>0.426</td>
<td>0.433</td>
<td>0.433</td>
<td>0.419</td>
<td>0.423</td>
<td>0.421</td>
</tr>
<tr>
<td></td>
<td>(0.836)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.893)</td>
<td>(0.531)</td>
<td>(0.669)</td>
</tr>
<tr>
<td>WMT</td>
<td>0.397</td>
<td>0.408</td>
<td>0.413</td>
<td>0.400</td>
<td>0.409</td>
<td>0.408</td>
</tr>
<tr>
<td></td>
<td>(0.882)</td>
<td>(0.026)</td>
<td>(0.008)</td>
<td>(0.690)</td>
<td>(0.036)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>XON</td>
<td>0.306</td>
<td>0.322</td>
<td>0.323</td>
<td>0.312</td>
<td>0.323</td>
<td>0.321</td>
</tr>
<tr>
<td></td>
<td>(0.882)</td>
<td>(0.001)</td>
<td>(0.065)</td>
<td>(0.110)</td>
<td>(0.000)</td>
<td>(0.005)</td>
</tr>
</tbody>
</table>
### Forecasting Results: All Series – Ten Days Ahead

<table>
<thead>
<tr>
<th>Series</th>
<th>STR-Tree/AE</th>
<th>STR-Tree/SB</th>
<th>STR-Tree/AE+SB</th>
<th>ARFIMA</th>
<th>HAR</th>
<th>EWMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>0.583</td>
<td>0.644</td>
<td>0.585</td>
<td>0.603</td>
<td>0.605</td>
<td>0.597</td>
</tr>
<tr>
<td></td>
<td>(0.927)</td>
<td>(0.003)</td>
<td>(0.865)</td>
<td>(0.064)</td>
<td>(0.175)</td>
<td>(0.368)</td>
</tr>
<tr>
<td>AIG</td>
<td>0.460</td>
<td>0.472</td>
<td>0.472</td>
<td>0.460</td>
<td>0.470</td>
<td>0.481</td>
</tr>
<tr>
<td></td>
<td>(0.856)</td>
<td>(0.249)</td>
<td>(0.238)</td>
<td>(0.874)</td>
<td>(0.469)</td>
<td>(0.122)</td>
</tr>
<tr>
<td>BA</td>
<td>0.512</td>
<td>0.551</td>
<td>0.530</td>
<td>0.523</td>
<td>0.515</td>
<td>0.533</td>
</tr>
<tr>
<td></td>
<td>(0.902)</td>
<td>(0.000)</td>
<td>(0.188)</td>
<td>(0.203)</td>
<td>(0.845)</td>
<td>(0.225)</td>
</tr>
<tr>
<td>CAT</td>
<td>0.499</td>
<td>0.541</td>
<td>0.520</td>
<td>0.507</td>
<td>0.517</td>
<td>0.513</td>
</tr>
<tr>
<td></td>
<td>(0.855)</td>
<td>(0.011)</td>
<td>(0.135)</td>
<td>(0.368)</td>
<td>(0.321)</td>
<td>(0.366)</td>
</tr>
<tr>
<td>GE</td>
<td>0.452</td>
<td>0.489</td>
<td>0.459</td>
<td>0.467</td>
<td>0.457</td>
<td>0.464</td>
</tr>
<tr>
<td></td>
<td>(0.884)</td>
<td>(0.000)</td>
<td>(0.616)</td>
<td>(0.015)</td>
<td>(0.776)</td>
<td>(0.415)</td>
</tr>
<tr>
<td>GM</td>
<td>0.455</td>
<td>0.526</td>
<td>0.469</td>
<td>0.481</td>
<td>0.486</td>
<td>0.483</td>
</tr>
<tr>
<td></td>
<td>(0.921)</td>
<td>(0.000)</td>
<td>(0.088)</td>
<td>(0.000)</td>
<td>(0.008)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>HP</td>
<td>0.744</td>
<td>0.759</td>
<td>0.756</td>
<td>0.746</td>
<td>0.745</td>
<td>0.731</td>
</tr>
<tr>
<td></td>
<td>(0.659)</td>
<td>(0.327)</td>
<td>(0.347)</td>
<td>(0.560)</td>
<td>(0.486)</td>
<td>(0.727)</td>
</tr>
<tr>
<td>INTC</td>
<td>0.607</td>
<td>0.754</td>
<td>0.640</td>
<td>0.629</td>
<td>0.640</td>
<td>0.625</td>
</tr>
<tr>
<td></td>
<td>(0.903)</td>
<td>(0.000)</td>
<td>(0.030)</td>
<td>(0.014)</td>
<td>(0.127)</td>
<td>(0.408)</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.456</td>
<td>0.477</td>
<td>0.485</td>
<td>0.460</td>
<td>0.468</td>
<td>0.487</td>
</tr>
<tr>
<td></td>
<td>(0.899)</td>
<td>(0.020)</td>
<td>(0.003)</td>
<td>(0.592)</td>
<td>(0.345)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>KO</td>
<td>0.411</td>
<td>0.448</td>
<td>0.433</td>
<td>0.414</td>
<td>0.429</td>
<td>0.430</td>
</tr>
<tr>
<td></td>
<td>(0.892)</td>
<td>(0.001)</td>
<td>(0.006)</td>
<td>(0.574)</td>
<td>(0.134)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>MRK</td>
<td>0.436</td>
<td>0.466</td>
<td>0.446</td>
<td>0.437</td>
<td>0.441</td>
<td>0.440</td>
</tr>
<tr>
<td></td>
<td>(0.891)</td>
<td>(0.000)</td>
<td>(0.189)</td>
<td>(0.753)</td>
<td>(0.625)</td>
<td>(0.541)</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.505</td>
<td>0.551</td>
<td>0.525</td>
<td>0.510</td>
<td>0.512</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>(0.862)</td>
<td>(0.000)</td>
<td>(0.081)</td>
<td>(0.250)</td>
<td>(0.727)</td>
<td>(0.354)</td>
</tr>
<tr>
<td>PFE</td>
<td>0.500</td>
<td>0.536</td>
<td>0.495</td>
<td>0.508</td>
<td>0.506</td>
<td>0.510</td>
</tr>
<tr>
<td></td>
<td>(0.540)</td>
<td>(0.000)</td>
<td>(0.938)</td>
<td>(0.180)</td>
<td>(0.243)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>WMT</td>
<td>0.524</td>
<td>0.536</td>
<td>0.527</td>
<td>0.518</td>
<td>0.519</td>
<td>0.511</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.001)</td>
<td>(0.296)</td>
<td>(0.734)</td>
<td>(0.535)</td>
<td>(0.689)</td>
</tr>
<tr>
<td>XON</td>
<td>0.395</td>
<td>0.410</td>
<td>0.400</td>
<td>0.396</td>
<td>0.427</td>
<td>0.432</td>
</tr>
<tr>
<td></td>
<td>(0.899)</td>
<td>(0.099)</td>
<td>(0.516)</td>
<td>(0.797)</td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>
### IBM Forecasting Results: Inclusion of Jumps

#### MAE, $R^2$ and Forecasting Accuracy Tests

<table>
<thead>
<tr>
<th></th>
<th>MAE</th>
<th>HLN</th>
<th>SPA</th>
<th>$R^2$</th>
<th>HLN</th>
<th>SPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STR-Tree/AE</td>
<td>0.324</td>
<td>0.000</td>
<td>0.340</td>
<td>0.644</td>
<td>0.001</td>
<td>0.785</td>
</tr>
<tr>
<td>HAR</td>
<td>0.334</td>
<td>0.079</td>
<td>0.001</td>
<td>0.621</td>
<td>0.259</td>
<td>0.004</td>
</tr>
<tr>
<td>5 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STR-Tree/AE</td>
<td>0.398</td>
<td>0.000</td>
<td>0.793</td>
<td>0.500</td>
<td>0.005</td>
<td>0.968</td>
</tr>
<tr>
<td>HAR</td>
<td>0.410</td>
<td>0.198</td>
<td>0.004</td>
<td>0.472</td>
<td>0.194</td>
<td>0.007</td>
</tr>
<tr>
<td>10 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STR-Tree/AE</td>
<td>0.450</td>
<td>0.008</td>
<td>0.504</td>
<td>0.386</td>
<td>0.068</td>
<td>0.742</td>
</tr>
<tr>
<td>HAR</td>
<td>0.463</td>
<td>0.480</td>
<td>0.014</td>
<td>0.355</td>
<td>0.033</td>
<td>0.041</td>
</tr>
<tr>
<td>20 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STR-Tree/AE</td>
<td>0.450</td>
<td>0.008</td>
<td>0.504</td>
<td>0.386</td>
<td>0.068</td>
<td>0.742</td>
</tr>
<tr>
<td>HAR</td>
<td>0.463</td>
<td>0.480</td>
<td>0.014</td>
<td>0.355</td>
<td>0.033</td>
<td>0.041</td>
</tr>
</tbody>
</table>
Conclusions

- We proposed a tree-structured multiple-regime model to describe the dynamics of the realized volatility of 16 DJIA stocks.
- The transitions between regimes were controlled by past cumulated returns.
- When put into proof in a forecasting exercise, the proposed model outperformed several linear and nonlinear alternatives, including the ARFIMA model.
We proposed a tree-structured multiple-regime model to describe the dynamics of the realized volatility of 16 DJIA stocks.

The transitions between regimes were controlled by past cumulated returns.

When put into proof in a forecasting exercise, the proposed model outperformed several linear and nonlinear alternatives, including the ARFIMA model.
We proposed a tree-structured multiple-regime model to describe the dynamics of the realized volatility of 16 DJIA stocks. The transitions between regimes were controlled by past cumulated returns. When put into proof in a forecasting exercise, the proposed model outperformed several linear and nonlinear alternatives, including the ARFIMA model.
We proposed a tree-structured multiple-regime model to describe the dynamics of the realized volatility of 16 DJIA stocks.

The transitions between regimes were controlled by past cumulated returns.

When put into proof in a forecasting exercise, the proposed model outperformed several linear and nonlinear alternatives, including the ARFIMA model.