# Regime Switching and Technical Trading with Dynamic Bayesian Networks in High-Frequency Stock Markets

*Luis Damiano, Brian Peterson, Michael Weylandt*

*2017-08-29*

## Contents

This work aims at replicating the Hierarchical Hidden Markov Model (HHMM) originally proposed by Tayal (2009) to learn price and volume patterns using rules from technical analysis, infer the hidden state of the system and identify runs and reversals out-of-sample in a statistically significant way. The main goal is to reproduce the results of the original research as well as to provide additional insight and criticism. Also, we produce public programming code in Stan (Carpenter et al. 2016) for a fully Bayesian estimation of the model parameters and inference on hidden quantities. A brief introduction about Hidden Markov Models (HMM) can be found in our literature review.

---

## 1 Motivation

The use of Markov-switching models to represent bull and bear markets is not novel. Chauvet and Potter (2000) define bulls and bears according to the general increasing or decreasing trend in prices and uses an unobservable two-state Markov variable to represent investors' real-time belief about the state of financial conditions. Maheu and McCurdy (2000) succeed in combining return mean and variance to identify all major market downturns in over 160 years of monthly price data. Lunde and Timmermann (2004), who define

market states based on sequences of stopping times tracing local peaks and troughs in prices, find duration dependence in stock prices. The longer a bull market has lasted, the lower is the probability that it will come to a termination. In contrast, the longer a bear market has lasted, the higher is its termination probability. Interest rates affect cumulated changes in stock prices as well. The existence of market states is not only supported by statistical models. Gordon and St-Amour (2000) establishes a link between time-varying prices of risk under the Capital Asset Pricing Model framework and state-dependent preferences.

# 2 Hierarchical Hidden Markov Models

The HHMM is a recursive hierarchical generalization of the HMM that provides a systematic unsupervised approach for complex multi-scale structure. The model is motivated by the multiplicity of length scales and the different stochastic levels (recursive nature) present in some sequences. Additionally, it infers correlated observations over long periods via higher levels of hierarchy.

The model structure is fairly general and allows an arbitrary number of activations of its submodels. The multi-resolution structure is handled by temporal experts[1] of different time scales.

## 2.1 Model specification

HHMM are structured multi-level stochastic processes that generalize HHM by making each of the hidden states an autonomous probabilistic model. There are two kinds of states: internal states are HHMM that emit sequences by a recursive activation of one of the substates, while production states generate an output symbol according to the probability distribution of the set of output symbols.

Hidden dynamics are lead by transitions. Vertical transitions involve the activation of a substate by an internal state, they may include further vertical transitions to lower level states. Once completed, they return the control to the state that originated the recursive activation chain. Then, a horizontal transition is performed. Its state transition within the same level.

A HHMM can be represented as a standard single level HMM whose states are the production states of the corresponding HHMM with a fully connected structure, i.e. there is a non-zero probability of transition from any state to any other state. This equivalent new model lacks the multi-level structure.

Let $z_t^d = i$ be the state of an HHMM at the step $t$, where $i \in \{1, \ldots, |z^d|\}$ is the state index, $|z^d|$ is the number of possible steps within the $d$-th level and $d \in \{1, \ldots, D\}$ is the hierarchy index taking values $d = 1$ for the root state, $d = \{2, \ldots, \ldots, D-1\}$ for the remaining internal states and $d = D$ for the production states.

In addition to its structure, the model is characterized by the state transition probability between the internal states and the output distribution of the production states. For each internal state $z_t^d$ for $d \in \{1, \ldots, D-1\}$, there is a state transition probability matrix $\boldsymbol{A}^d$ with elements $a_{ij}^d = p(z_t^{d+1} = j|z_t^{d+1} = j)$ being the probability of a horizontal transition from the $i$-th state to the $j$-th state within the level $d$. Similarly, there is the initial distribution vector over the substates $\boldsymbol{\pi}^d$ with elements $\pi_j^d = p(z_t^{d+1} = j|z_t^d)$ for $d \in \{1, \ldots, D-1\}$. Finally, each production state $z_t^D$ is parametrized by the output parameter vector $\boldsymbol{\theta}_o^i$ whose form depends on the specification of the observation model $p(\boldsymbol{x}_t|z_t^D = j, \boldsymbol{\theta}_o^j)$ corresponding to the $j$-th production state.

## 2.2 Generative model

The root node initiates a stochastic sequence generation. An observation for the first step in the sequence $t$ is generated by drawing at random one of the possible substates according to the initial state distribution $\boldsymbol{\pi}^1$. To replicate the recursive activation process, for each internal state entered $z_t^d$ one of the substates

---

[1]In Machine Learning terminology, a problem is divided into homogeneous regions addressed by an expert submodel. A gating network or function decides which expert to use for each input region.

is randomly chosen according to the corresponding initial probability vector $\boldsymbol{\pi}^d$. When an internal state transitions to a production state $z_t^D = j$, a single observation is generated according to the state output parameter vector $\boldsymbol{\theta}_o^j$. Control returns to the internal state that lead to the current production state $z_t^{D-1}$, which in turns selects the next state in the same level according to transition matrix $\boldsymbol{A}^{D-1}$.

Save for the top, each level $d \in \{2, \ldots, D\}$ has a final state that terminates the stochastic state activation process and returns the control to the parent state of the whole hierarchy. The generation of the observation sequence is completed when control of all the recursive activations returns to the root state.

## 2.3 Parameter estimation

The parameters of the models are $\boldsymbol{\theta} = \left\{ \left\{ \boldsymbol{A}^d \right\}_{d \in \{1, \ldots, D-1\}}, \left\{ \boldsymbol{\pi}^d \right\}_{d \in \{1, \ldots, D-1\}}, \left\{ \boldsymbol{\theta}_o \right\} \right\}$. The form of $\boldsymbol{\theta}_o$ depends on the specification of the production states. We refer the reader to Fine and Singer (1998) for a detailed treatment of estimation and inference procedures.

---

# 3 Regime Switching and Technical Trading with Dynamic Bayesian Networks in High-Frequency Stock Markets

## 3.1 Preamble

Many published works argue for the existence of systematic patterns in price action. Brock, Lakonishok, and LeBaron (1992) test technical trading rules and strategies. Karpoff (1987), later followed by Gallant, Rossi, and Tauchen (1992), make four contributions supporting that volume is a significant source of information for price changes. Lo, Mamaysky, and Wang (2000) argue that several technical indicators provide incremental information and may have some practical value. Statistical significance does not imply profitability, however, and part of the research towards technical analysis is subject to various problems in their testing procedure (Park and Irwin 2007).

Tayal (2009) is first to propose a graphical model for technical analysis. The author creates data features based on technical analysis rules and designs a HHMM to learn intraday price and volume patterns. Probabilistic inference allows the identification of two distinct states in high-frequency data that are mainly marked by buying and selling pressure.

## 3.2 Feature extraction

### 3.2.1 Input series

Tick series are a sequence of triples $\{y_k\}$ with $y_k = (t_k, p_k, v_k)$, where $t_k \leq t_{k+1}$ is the time stamp in seconds, $p_k$ is the trade price and $v_k$ is the trade volume. The sequence is ordered by the occurrence of trades. There can be more than one trade within a second.

Following Tayal (2009), who in turns drew inspiration from the technical analysis techniques proposed by Ord (2008), we derive a zig-zag sequence that captures the bid-ask bounce $\{z_k\}$ with $z_k = (i_n, j_n, e_n, \phi_n)$, where $i_n \leq i_j$ are indices to the tick series representing the starting and ending point of the extrema, $e_n = p_k \; \forall \; k : i_n \leq k \leq j_n$ is the price at the local extrema, and $\phi_m$ measures the average volume per second during the zig-zag leg ending at $e_n$:

$$\phi_n = \frac{1}{t_{j_n} - t_{i_{n-1}} + 1} \sum_{k=i_{n-1}}^{j_n} v_k.$$

We note that $p_{i_n} < e_n < p_{j_n+1}$ for local maxima and $p_{i_n} > e_n > p_{j_n+1}$ for local minima. The average volume, which includes the end-point extrema, is normalized by $t_{j_n} - t_{i_{n-1}} + 1$ to avoid division by zero when the zig-zag leg occurs within the same time period. Most importantly, we underline that the $n$-th zig-zag point $z_n$ is realized only after observing the $(j_n + 1)$-th tick point $y_{j_n+1}$. Failing to consider the one tick lag between leg completion and the time of detection would cause look-ahead bias in the out of sample forecasts.

### 3.2.2   Processing rules

Discrete features are created based on the zig-zag series $\{z_n\}$. We first create an auxiliary series $\{O_n\}$ with $O_n = (f_n^0, f_n^1, f_n^2)$, where $f_n^0$ represents the direction of the zig-zag, $f_n^1$ indicates the existence of a trend and $f_n^2$ indicates whether average volume increased or decreased.

Formally,

$$f_n^0 = \begin{cases} +1 & \text{if } e_n \text{ is a local maximum (positive zig-zag leg)} \\ -1 & \text{if } e_n \text{ is a local minimum (negative zig-zag leg)}, \end{cases}$$

and

$$f_n^1 = \begin{cases} +1 & \text{if } e_{n-4} < e_{n-2} < e_n \wedge e_{n-3} < e_{n-1} \text{ (up-trend)} \\ -1 & \text{if } e_{n-4} > e_{n-2} > e_n \wedge e_{n-3} > e_{n-1} \text{ (down-trend)} \\ 0 & \text{otherwise (no trend).} \end{cases}$$

For the third indicator function, we compute the average volume ratios,

$$\nu_n^1 = \frac{\phi_n}{\phi_{n-1}}, \quad \nu_n^2 = \frac{\phi_n}{\phi_{n-2}}, \quad \nu_n^3 = \frac{\phi_{n-1}}{\phi_{n-2}},$$

we transform the ratios into a discrete variable using an arbitrary threshold $\alpha$,

$$\tilde{\nu}_n^j = \begin{cases} +1 & \text{if } \nu_n^j - 1 > \alpha \\ -1 & \text{if } 1 - \nu_n^j > \alpha \\ 0 & \text{if } |\nu_n^j - 1| \le \alpha, \end{cases}$$

and we finally define

$$f_n^2 = \begin{cases} +1 & \text{if } \tilde{\nu}_n^1 = 1, \tilde{\nu}_n^2 > -1, \tilde{\nu}_n^3 < 1 \text{ (volume strengthens)} \\ -1 & \text{if } \tilde{\nu}_n^1 = -1, \tilde{\nu}_n^2 < -1, \tilde{\nu}_n^3 > -1 \text{ (volume weakens)} \\ 0 & \text{otherwise (volume is indeterminant).} \end{cases}$$

The features or legs $\boldsymbol{D} = \{D_1, \ldots, D_9\}$, $\boldsymbol{U} = \{U_1, \ldots, U_9\}$ are then created using the Table 1.

Defining appropriate rules that capture trade volume dynamics and identify trends in volume despite high-frequency noise is the most challenging aspect of this design. Wisebourt (2011) proposes a modification for the feature extraction procedure. By computing the spread between the Volume Weighted Average Prices of the bid and the ask, he devices a book imbalance metric that describes the state of the order book at any given moment in time. In turns, Sandoval and Hernández (2015) applies wavelets over two simple-smoothed exponential distance-weighted average volume series to measure trade volume concentration in both sides of the book.

|       | Zig-zag direction | Price trend | Change in volume | Market State |
|-------|-------------------|-------------|------------------|--------------|
| $D_1$ | Down -1 | Up +1 | Weak -1 | Bull |
| $D_2$ | Down -1 | Down -1 | Weak -1 | Bull |
| $D_3$ | Down -1 | Up +1 | Intederminant 0 | Bull |
| $D_4$ | Down -1 | No trend 0 | Weak -1 | Bull |
| $D_5$ | Down -1 | No trend 0 | Intederminant 0 | Local volatility |
| $D_6$ | Down -1 | No trend 0 | Strong +1 | Bear |
| $D_7$ | Down -1 | Down -1 | Intederminant 0 | Bear |
| $D_8$ | Down -1 | Up +1 | Strong +1 | Bear |
| $D_9$ | Down -1 | Down -1 | Strong +1 | Bear |
| $U_1$ | Up +1 | Up +1 | Strong +1 | Bull |
| $U_2$ | Up +1 | Down -1 | Strong +1 | Bull |
| $U_3$ | Up +1 | Up +1 | Intederminant 0 | Bull |
| $U_4$ | Up +1 | No trend 0 | Strong +1 | Bull |
| $U_5$ | Up +1 | No trend 0 | Intederminant 0 | Local volatility |
| $U_6$ | Up +1 | No trend 0 | Weak -1 | Bear |
| $U_7$ | Up +1 | Down -1 | Intederminant 0 | Bear |
| $U_8$ | Up +1 | Up +1 | Weak -1 | Bear |
| $U_9$ | Up +1 | Down -1 | Weak -1 | Bear |

Table 1: Feature space.

## 3.3 Model

We adhere to the methodology proposed in the original work as much as possible. We set up a HHMM to learn the sequence of discrete features extracted from a high-frequency time series of stock prices and traded volume. The figure below summarizes the model structure in the form of a Dynamic Bayesian Network.

The graph starts with a root node $z^0$ that has two top-level children $z_1^1$ and $z_2^1$ representing bullish markets (or runs) and bearish markets (or reversals). The specifications do not pose any constraints that determines beforehand which node takes each of the possible two meanings. In consequence, latent states need to be labeled after the learning stage based on sample characteristics such as mean returns. Although the original author does not mention this possibility, prior information, like parameter ordering, could be embedded to break symmetry and mitigate eventual identification issues.
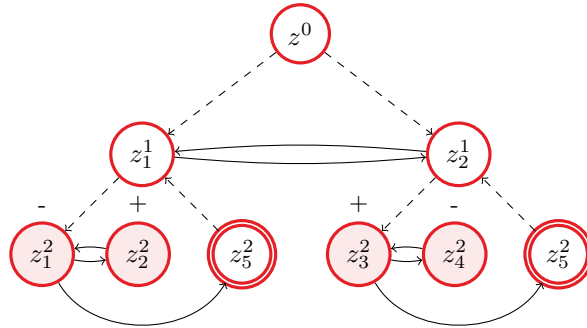


Figure 1: Hierarchical Hidden Markov Model for price and volume.

Each top node activates a probabilistic HMM with two latent states for negative and positive zig-zag legs. The sub-model activated by the node $z_1^1$ always starts with node $z_1^2$ producing an observation from the distribution of negative zig-zag $\boldsymbol{D} = \{D_1, \dots, D_9\}$. Next, it transitions to (a) node $z_2^2$ producing a positive zig-zag leg $\boldsymbol{U} = \{U_1, \dots, U_9\}$, or (b) the end node and switches to the second sub-model, landing on node $z_3^2$ and producing a positive zig-zag leg. Restricting transitions to this limited set of movements force alternation between positive and negative zig-zag legs, thus guaranteeing that all possible observation sequences are well behaved. The sub-model belonging to $z_2^1$ has similar but symmetrical behaviour. These two inner models are conditionally independent, an advantage we will take for computational time.

The persistence of price trends vary according to the resolution of the dataset. In low frequency contexts, financial analysts will frequently define short and long term trends based on timespans involving weeks and months of observations. In high frequency trading, however, trends may last from seconds to hours. The main idea behind the conception of the hierarchical model is that top levels may group a sequence of price movements forming a general trend, whereas the bottom nodes allow zig-zags to represent micro trends that may deviate shortly from the main drift. That is, the model accommodates for possibly unequally probable negative price movements in an upwards market as well as positive short trends during a price reversal.

Theoretically, all HHMM can be expressed as an equivalent HMM with possibly sparse initial probability vector and transition matrix. Although learning from this representation may prove less efficient in terms of computational complexity, the relatively simple structure and many restrictions of the model under study make estimation and inference feasible. The equivalent "expanded" HMM has $K = 4$ states, the following $K$-sized initial probability vector

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_1 & 0 & 1 - \pi_1 & 0 \end{bmatrix}$$

and the $K \times K$ transition matrix

$$\boldsymbol{A} = \begin{bmatrix} 0 & a_{12} & 1 - a_{12} & 0 \\ 1 & 0 & 0 & 0 \\ a_{31} & 0 & 0 & 1 - a_{31} \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

where each element $a_{ij}$ represents the probability of transitioning from the hidden state $i$ (row) to $j$ (column) in one step. The matrix is sparse with zeros representing nodes with no direct connections. Since the initial probability vector and the rows of the transition matrix must sum to one, hidden dynamics are governed by only three free parameters.

Production nodes emit one observation from the finite sets of possible outputs $\boldsymbol{O}_D$ (for $z_1^2$ and $z_4^2$) and $\boldsymbol{O}_U$ (for $z_2^2$ and $z_3^2$). Conditional probability distributions are given by output probability vectors of length $L = 9$ subject to sum-to-one constraints:

$$\begin{aligned} \boldsymbol{B}^1 = p(\boldsymbol{x}_t | z_1^2) &= \left[ b_{D_1}^1, \ldots, b_{D_9}^1 \right], \\ \boldsymbol{B}^2 = p(\boldsymbol{x}_t | z_2^2) &= \left[ b_{U_1}^2, \ldots, b_{U_9}^2 \right], \\ \boldsymbol{B}^3 = p(\boldsymbol{x}_t | z_3^2) &= \left[ b_{U_1}^3, \ldots, b_{U_9}^3 \right], \\ \boldsymbol{B}^4 = p(\boldsymbol{x}_t | z_4^2) &= \left[ b_{D_1}^4, \ldots, b_{D_9}^4 \right]. \end{aligned}$$

The observation model has 32 free parameters.

On the whole, the parameter vector for the HMM representation reduces to vector of size 35,

$$\boldsymbol{\theta} = (\pi_1, a_{12}, a_{31}, b_{D_l}^1, b_{U_l}^2, b_{U_l}^3, b_{D_l}^4),$$

with $l \in \{1, \ldots, L - 1\}$.


## 3.4   Dataset

The original work presents results for both simulated and real data. The latter is based on historical high-frequency time series for the 60 stocks listed in the S&P/TSE60 index. The dataset consists of all 22 business days of May 2007. The author excludes three days due to significant errors without disclosing the exact dates. We confirm that our results are consistent with the original work by focusing on GoldCorp Inc (TSE:G), the only series described exhaustively in the original work.

Note that we do not model prices directly. Instead, non-linear transformations are applied to the trade price and volume series to produce the sequence of features that feeds the proposed model.

## 3.5 Methodology

Model parameters are estimated on a rolling window with five days each. Since top nodes are symmetrical, states are labeled ex-post based on the order of the in-sample mean of the percentage change in the initial and final price before the top level state switch. The state with larger and smaller returns are marked as bullish (a run) and bearish (a reversal) respectively.

After learning and labeling, the author runs two out of sample inference procedures on the sixth day. First, offline smoothing infers the hidden state at time $t$ based on the full evidence of the sixth day. Although this quantity is not useful for trading because of its look-ahead bias, it provides an upper bound benchmark for the model. Second, online filtering is used as a trading rule. Although smoothing is a valid benchmark, we focus exclusively on the latter.

Most of the diagnostics are based on trade returns. For the $l$-th top-level state switch, the percentage return is defined as

$$R_l = \frac{p_l^e - p_l^s}{p_l^s},$$

where $p_l^s$ and $p_l^e$ are the price at the start and end of the switch.

The information content of learning the top-level state is assessed by comparing the unconditional empirical distribution of trade returns versus their empirical conditional distribution given the top state. Additionally, regime return characteristics are validated: mean trade returns are expected to be higher for bullish regimes compared to bearish regimes, and they are expected to be positive and negative for runs and reversals respectively. In the original work, most of these analysis are run both in-sample and out-of-sample. We focus on out-of-sample results only.

Finally, a trading strategy is tested. After a zig-zag leg is completed, the trading system buys one unit every time the top-level state switches to bullish and sells one unit every time it switches to bearish. As an addition to the original research, where trades are executed one tick after the leg is observed to ensure that there is no lock-ahead bias, we investigate the decay of the strategy performance for longer lags.

## 3.6 GoldCorp Inc (TSE:G)

We present an in-depth study of one stock to assess the strengths and weaknesses of the model. Using the tick-by-tick series of GoldCorp Inc (TSE:G), we split our dataset in training (2007-05-04 to 2007-05-10 - five trading days) and test (2007-05-11) sets. Next, we run our procedure in a walking forward fashion for this stock as well as others and present some summary statistics for the performance of the strategy.

### 3.6.1 Data exploration

We center our attention on the sequence of trades, disregarding possibly valuable information from the bid and ask series. Future research may employ such information to improve model predictability. We start by extracting the features using the procedure detailed above. We set the threshold for the change in volume indicator variable in $\alpha = 0.25$ as suggested by the author of the original work. In-sample dataset reduced to 8386 zig-zags.

Apart from the zig-zags themselves, local extrema are interesting on their own. Although we could not gain insight by visually inspecting other intermediate indicators such as the trend $f_1$, further research may find value in them.
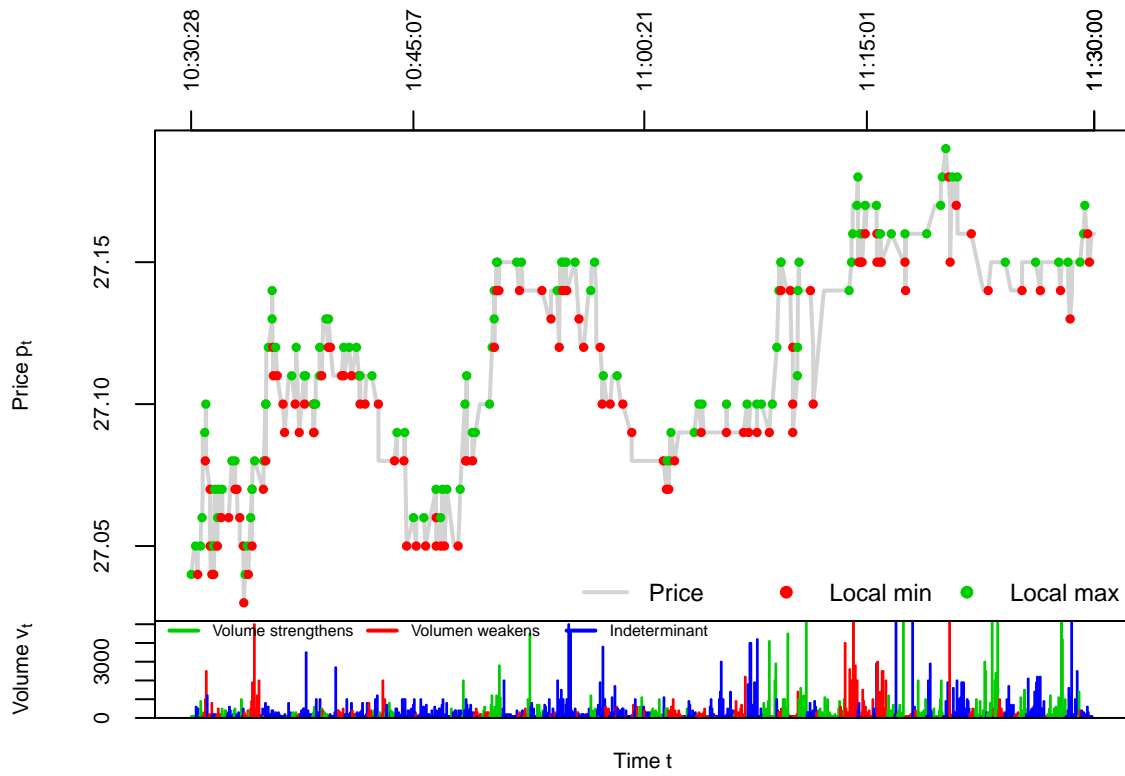
Figure 2: Local extrema detected in TSE:G 2007-05-11 10:30:00/2007-05-11 11:30:00.

The Features
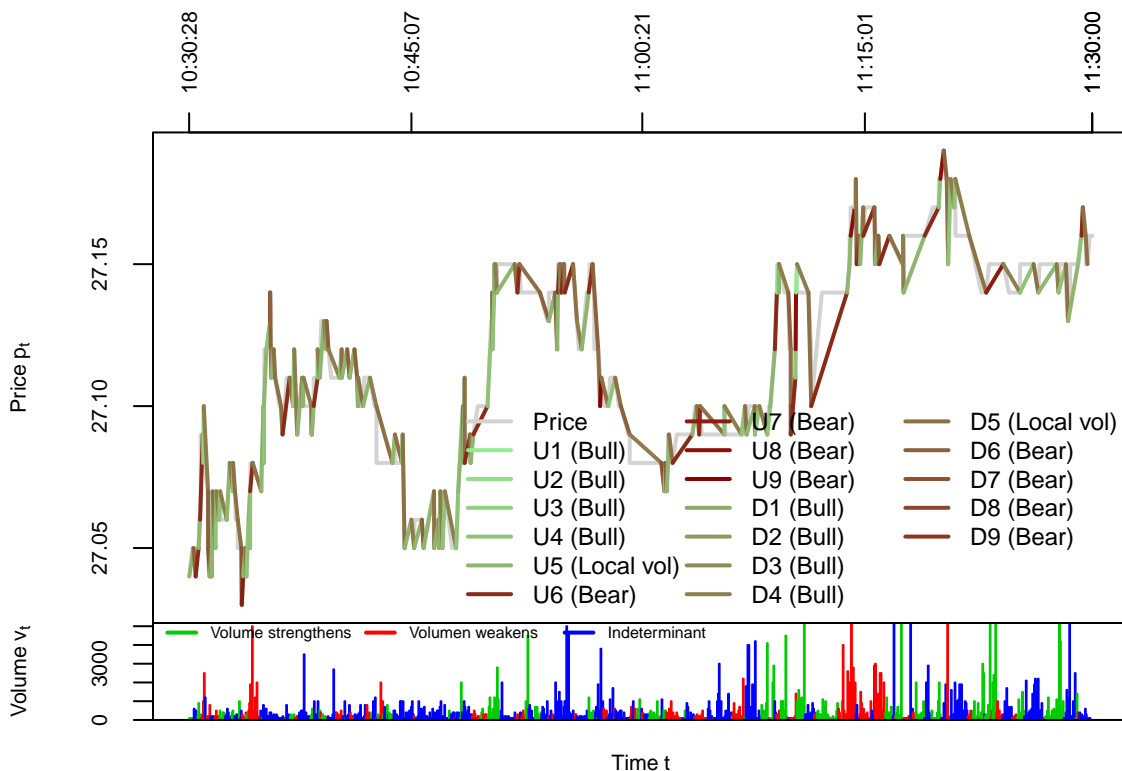
Figure 3: Features extracted from TSE:G 2007-05-11 10:30:00/2007-05-11 11:30:00.

### 3.6.2 Estimation

Model parameters can be segregated into two groups: parameters for the hidden model, including the initial distribution vector as well as the transition matrix, and parameters for the conditional multinomial distributions for the output. Table 2 summarizes the former. The estimates for the initial distribution probabilities $\pi_1$ and $1 - \pi_1$ are uncertain to a large extent, which is unsurprising as the sample provides with only one starting observation and the model imposes almost no prior information for the parameters. Team (2017), a technical manual for a programming language that also contains many brief discussion of statistical notions, describes some alternative specifications. If the sample were conceived as a subsequence of a long-running data generating process, the initial probabilities may be set to equal the stationary state probabilities of the transition Markov chain. Contrarily, if the sample was considered a finite-length sequence, the model may have a different starting distribution.

On the other hand, estimation of the transition parameters profit from a larger number of trades: more information reduce uncertainty. We note that the bull top state seems persistent as positive zig-zags $z_3^2$ are more likely to transition to bull negative zig-zags $z_4^2$ versus bear negative zig-zags $z_1^2$ by a factor of $a_{34}/a_{31} \approx 10$. After studying several stocks and timespans, we decide that these empirical observations are specific to the sample and are highly influenced by the general price trend present in the five-day training dataset.

Output distributions are summarized in Figure 4. It is important to note that zig-zags with no price trends $D_4, D_5, D_6, U_4, U_5, U_6$ are the most preeminent features. In particular, it is worthwhile to analyse the behaviour of the model in the presence of local volatility $D_5, U_5$, i.e. when there is no clear trends in price

9

|  | Mean | Std. Deviation | $q_{10\%}$ | $q_{50\%}$ | $q_{90\%}$ |
|---|---|---|---|---|---|
| $\pi_1$ | 0.51 | 0.28 | 0.10 | 0.52 | 0.89 |
| $1 - \pi_1$ | 0.49 | 0.28 | 0.11 | 0.48 | 0.90 |
| $a_{12}$ | 0.46 | 0.11 | 0.31 | 0.47 | 0.58 |
| $1 - a_{12}$ | 0.54 | 0.11 | 0.42 | 0.53 | 0.69 |
| $a_{21}$ | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 |
| $a_{31}$ | 0.09 | 0.06 | 0.02 | 0.08 | 0.17 |
| $1 - a_{31}$ | 0.91 | 0.06 | 0.83 | 0.92 | 0.98 |
| $a_{43}$ | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 |

Table 2: Estimated parameters of the transition matrix for TSE:G 2007-05-04 09:30:00/2007-05-10 16:30:00.

and volume.
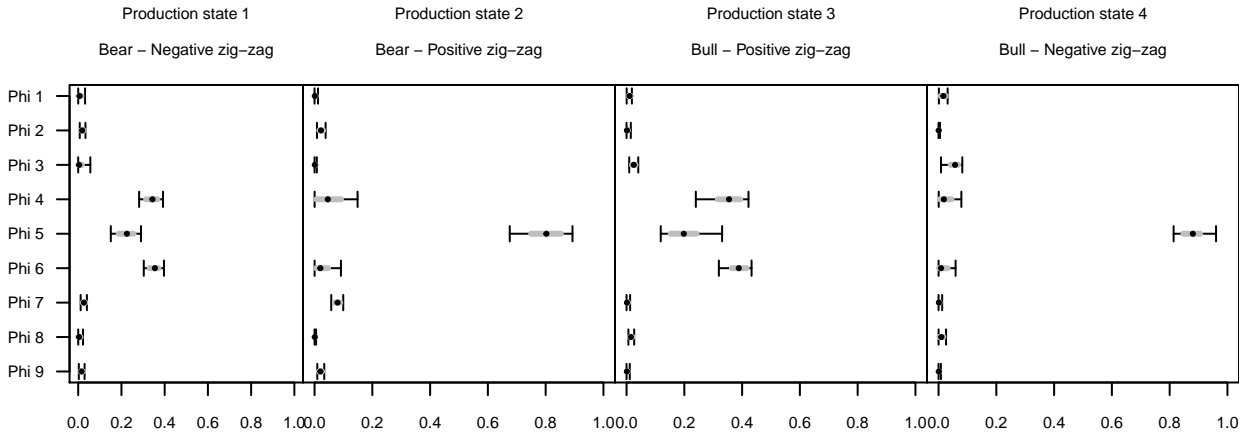


Figure 4: Estimated parameters of the conditional multinomial distribution of the outputs given the emission state (bottom node) for TSE:G 2007-05-04 09:30:00/2007-05-10 16:30:00.

We first study short-term price changes in the opposite direction to the long-term trend. The vast majority of negative zig-zags observed in bullish markets are due to local volatility ($\hat{\phi}_{45} = 0.88$) as do most of the positive zig-zags found in bearish markets ($\hat{\phi}_{25} = 0.80$). We note, however, that negative legs with no price trends but weakening or strengthening trade volume are equally probable in bearish markets ($\hat{\phi}_{14} \approx \hat{\phi}_{16}$). Analogously, we find in bullish markets that $\hat{\phi}_{34} \approx \hat{\phi}_{36}$. These observations counter the a priori classification stated in Table 1.

Moreover, we find that positive zig-zags originated in local volatility are more likely to be seen in bearish markets versus bullish markets by a factor of $\phi_{25}/\phi_{35} \approx 4$. The odds for negative zig-zags are similar. We warn the reader that the current model needs more information to classify this feature. All in all, we are warned that zig-zag direction $f^0$ and change in volume $f^2$ are not decisive without a price trend, a hint that the current model needs either better feature engineering rules for the change in volume or more external information to deal with local volatility.

### 3.6.3 Convergence

Figure 5 illustrates the trace plot of some arbitrary parameters as well as some diagnostic measures. In general terms, mixing and convergence to the stationary distribution is acceptable. The shrink factor of Gelman and Rubin (1992), close to 1 for all hidden quantities, suggest an adequate degree of convergence.

Sampling is efficient as signaled by effective sample size ratios near to 1 and Monte Carlo Standard Error to posterior standard deviation ratios well below 10%.
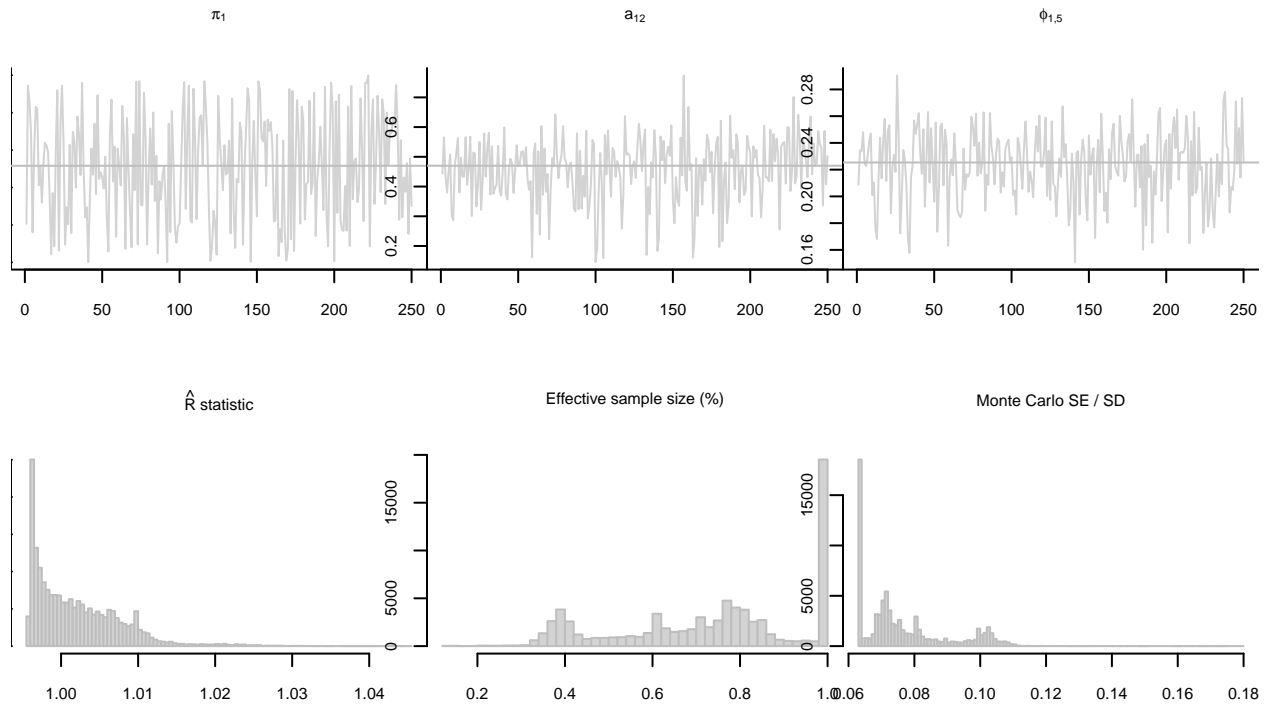


Figure 5: Traceplot of some arbitrarily selected parameters and histograms of diagnostic measures. Mixing, convergence to the stationary distribution and sampling efficiency are acceptable.

Although reestating the original Hierarchical Hidden Markov Model into a Hidden Markov Model may increase time and memory complexity, the new model becomes significantly easier to program and convergence. A full list of our results, including convergence statistics such as $\hat{R}$ and the effective sample size, is included in the Appendix Section 5.1.

### 3.6.4   State probability

The forward algorithm allows us to calculate the filtered belief state: the probability that an observation at time $t$ was emitted by one of the possible four states (bottom-nodes) given the evidence available up to $t$. We assign each observation to the emission state with largest filtered probability and, by the definition of the hierarchical model, they become naturally linked to one of the two possible top states (bears and bulls). Figure 6 reflects the resulting classification.

11
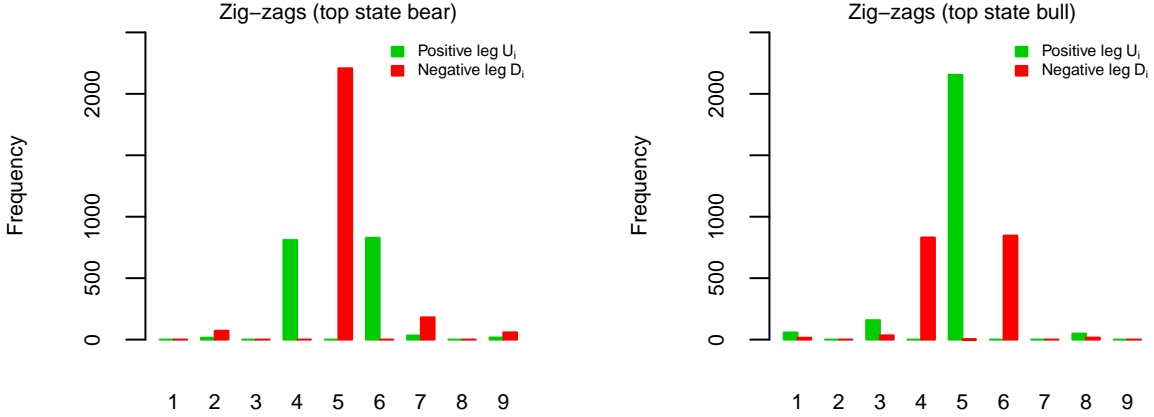
Figure 6: Distribution of features conditional on the estimated hidden regime (top node).

The following table provides some summary statistics for the returns of the observations classified in each state. The structure of the top nodes are symmetrical and they do not have an a priori order. We label them according to the in-sample mean trade returns.

| | Mean | SD | Skewness | Kurtosis | $q_{25}\%$ | $q_{50}\%$ | $q_{75}\%$ | Mean length | Median length |
|---|---|---|---|---|---|---|---|---|---|
| Bear | -0.01 | 0.06 | 7.97 | 237.46 | -0.04 | 0.00 | 0.00 | 10.62 | 6.00 |
| Bull | 0.00 | 0.05 | 0.97 | 9.37 | 0.00 | 0.00 | 0.04 | 10.17 | 6.00 |
| Unconditional | -0.00 | 0.06 | 5.53 | 173.81 | -0.04 | 0.00 | 0.00 | 10.40 | 6.00 |

Table 3: Summary statistics for the return of the trades assigned to each of the two possible top states for TSE:G 2007-05-04 09:30:00/2007-05-10 16:30:00. Trade returns are computed as defined in Section 3.5. Trade length is computed as the number of ticks involved. Returns expressed in percentage. SD means Standard Deviation.

Bull top states have a greater mean return than bear top state by construction. As it is also visible in Figure 7, a positive skewness coefficient for all states indicates that negative returns tend to overweight positive returns for this specific stock during the five-day in-sample dataset. However, bear markets have a marked skew towards negative returns and become more risky in term of extreme events (higher kurtosis).
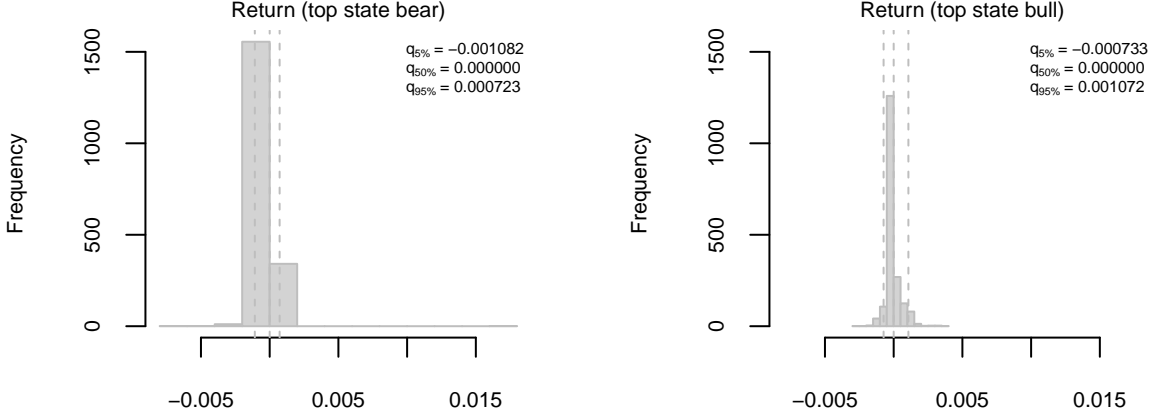
Figure 7: Conditional distribution of trade returns given the estimated top state in TSE:G 2007-05-04 09:30:00/2007-05-10 16:30:00.

We remark that trade returns originated in different stocks or timespans do not necessarily share these characteristics. Location, dispersion, symmetry and shape of the return distribution vary along stocks and day of analysis.

### 3.6.5 Fitted output

As mentioned in Section 3.3, the model assumes that outputs are emitted by one of four possible bottom nodes. By definition, negative legs belong to either $z_1^2$ or $z_4^2$ while positive legs belong to $z_2^2$ or $z_3^2$. Once the parameters are estimated, states $z_1^2$ and $z_2^2$ are labeled as bears while $z_3^2$ and $z_4^2$ are marked as bulls according to the mean trade return of the observations belonging to each top node.

Bullish states allow for negative zig-zags and bearish states allow for positive zig-zags as long as the trade volume is indeterminant or weak. The results of the classification are summarized in Table 4. As expected, the bullish top-node capture positive movements due to local volatility as well as downward movements with weak volume.

We remark that all observations belonging to each of the 18 possible features are imputed to one state. For example, observations $U_3$ are all mapped into state $z_4^2$. In a sense, hard classification suffers from information loss. This requires further research as may be seen as a weakness of the inference procedure. Nonetheless, in-sample classification results do not look alarming in Figure 9.

| Top | Bottom | $U_1$ | $U_2$ | $U_3$ | $U_4$ | $U_5$ | $U_6$ | $U_7$ | $U_8$ | $U_9$ | $D_1$ | $D_2$ | $D_3$ | $D_4$ | $D_5$ | $D_6$ | $D_7$ | $D_8$ | $D_9$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bear | 1 | 0 | 15 | 0 | 810 | 0 | 828 | 33 | 0 | 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bear | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 72 | 0 | 0 | 2208 | 0 | 181 | 0 | 59 |
| Bull | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 0 | 15 | 0 | 34 | 831 | 0 | 846 | 0 | 16 | 0 |
| Bull | 4 | 58 | 0 | 158 | 0 | 2155 | 0 | 0 | 22 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

Table 4: Features extracted are classified as emitted by one of the four possible bottom-nodes according to the filtered probability. TSE:G 2007-05-04 09:30:00/2007-05-10 16:30:00.

Now, look how at how the features are classified: Mention plot number on tayal.
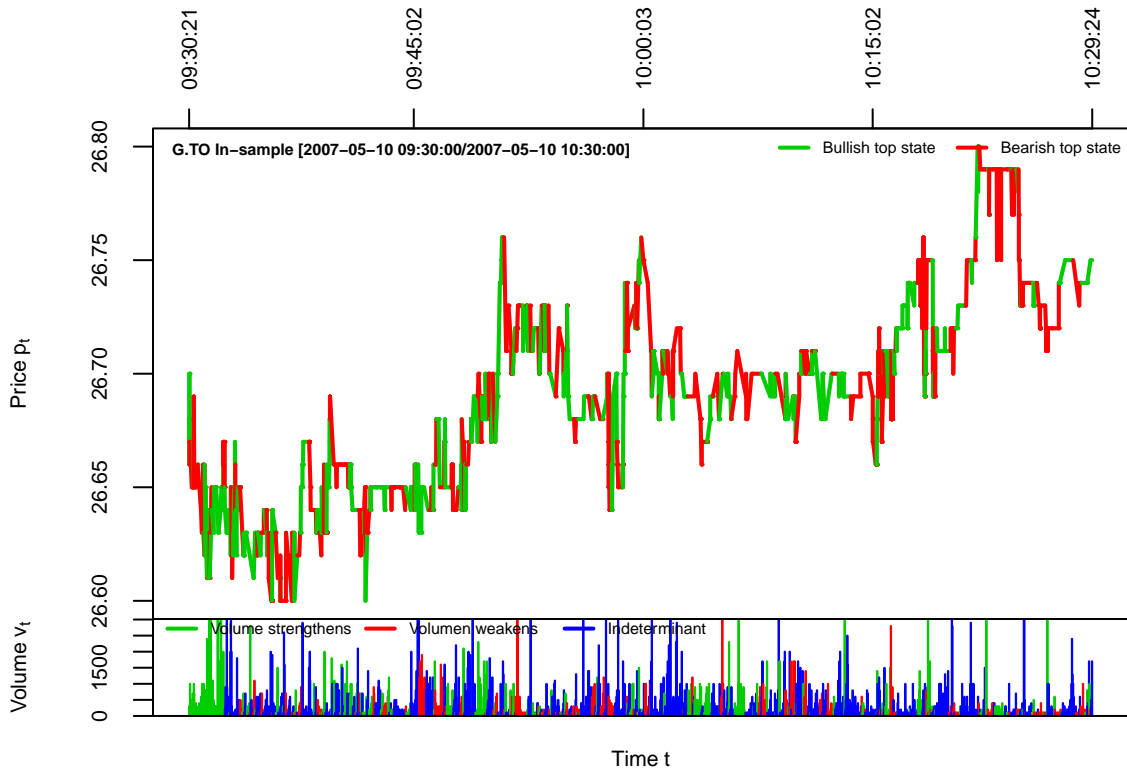
Figure 8: Tick by tick sequence of trades, classified as belonging to the bear or the bull top state, from TSE:G 2007-05-10 09:30:00/2007-05-10 10:30:00.

### 3.6.6 Trading Strategy

Using the parameters estimated from the five days of training data, we run the trading strategy out of sample during one day. Figure 9 depicts the equity line for 2007-05-10 while Table 5 summarises the returns for a trading month. The HHMM strategy can produce a positive result in a day with a downward price trend. Positive returns can be achieved in the presence of execution lag, even though trading results are highly sensitive. In many cases, the strategy does not only contribute to the generation of profits but also to risk reduction.

Figure 9: Out of sample equity line for TSE:G 2007-05-10 09:30:00/2007-05-10 10:30:00.

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | -1.24 | 3.99 | -1.52 | -0.92 | 0.70 | 0.62 | 1.74 |
| 2007-05-09 | -0.41 | 3.93 | 0.33 | 1.82 | 1.89 | 0.55 | 0.77 |
| 2007-05-10 | -0.37 | 4.19 | 1.18 | -0.61 | 1.81 | 1.73 | 0.18 |
| 2007-05-11 | -0.04 | 0.18 | 0.10 | 1.13 | -0.50 | -0.64 | 0.29 |
| 2007-05-14 | -3.33 | 2.71 | -1.33 | 0.63 | -0.95 | -0.46 | -1.20 |
| 2007-05-15 | -0.04 | 3.48 | -0.16 | 0.06 | 1.83 | 2.06 | 0.12 |
| 2007-05-16 | -0.42 | 5.45 | -0.78 | -0.38 | 1.23 | 2.80 | -2.38 |
| 2007-05-17 | -0.12 | -1.78 | 0.09 | 2.41 | 0.42 | -2.97 | -0.25 |
| 2007-05-18 | 1.25 | -1.02 | 0.70 | 0.38 | 2.20 | 1.41 | 1.73 |
| 2007-05-22 | -2.39 | -1.92 | -1.89 | 1.70 | 0.52 | 2.39 | 2.16 |
| 2007-05-23 | -1.02 | 1.72 | -0.11 | -0.65 | -0.73 | 0.96 | 1.45 |
| 2007-05-24 | -3.18 | 2.45 | -0.25 | -0.92 | -0.74 | -0.00 | -1.61 |
| 2007-05-25 | 0.33 | -1.36 | -1.44 | 0.06 | 0.69 | -1.83 | -1.72 |
| 2007-05-28 | -0.81 | -1.79 | -0.90 | 0.65 | -1.42 | -1.01 | 1.51 |
| 2007-05-29 | -2.25 | -1.53 | -1.30 | -1.80 | -0.10 | 2.12 | 0.97 |
| 2007-05-30 | 1.41 | -2.21 | -3.49 | -3.10 | -1.25 | -3.88 | -2.30 |
| 2007-05-31 | 3.96 | 0.20 | -1.32 | -0.86 | -1.70 | -1.33 | -2.97 |
| Total | -8.56 | 17.44 | -11.51 | -0.54 | 3.84 | 2.25 | -1.71 |
| Min | -3.33 | -2.21 | -3.49 | -3.10 | -1.70 | -3.88 | -2.97 |
| Mean | -0.51 | 0.98 | -0.71 | -0.02 | 0.23 | 0.15 | -0.09 |
| Median | -0.41 | 0.20 | -0.78 | 0.06 | 0.42 | 0.55 | 0.18 |
| Max | 3.96 | 5.45 | 1.18 | 2.41 | 2.20 | 2.80 | 2.16 |
| SD | 1.78 | 2.63 | 1.13 | 1.38 | 1.27 | 1.92 | 1.65 |

Table 5: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). TSE:G.

### 3.6.7 Application to other stocks

We run the walk forward backtesting procedure for twelve other stocks: BBDb.TO, BCE.TO, CTCa.TO, ECA.TO, G.TO, K.TO, MGa.TO, NXY.TO, SJRb.TO, SU.TO, TCKb.TO, TLM.TO. Using 17 days of data, we train the model on a five-day rolling window and then apply it to trade out of sample during a day. We collect the returns from seven configurations, including the buy & hold and the HHMM strategies with none to five ticks of lags (we remind the reader that no lag implies look-ahead bias). In total, we compute $12 \times 17 \times 7 = 1,428$ daily returns. As the correlation matrix in Table 6 shows, the HHMM strategy is virtually uncorrelated with buy and hold. Section 5.2 in the appendix details the trade returns and equity line for the backtested stocks.
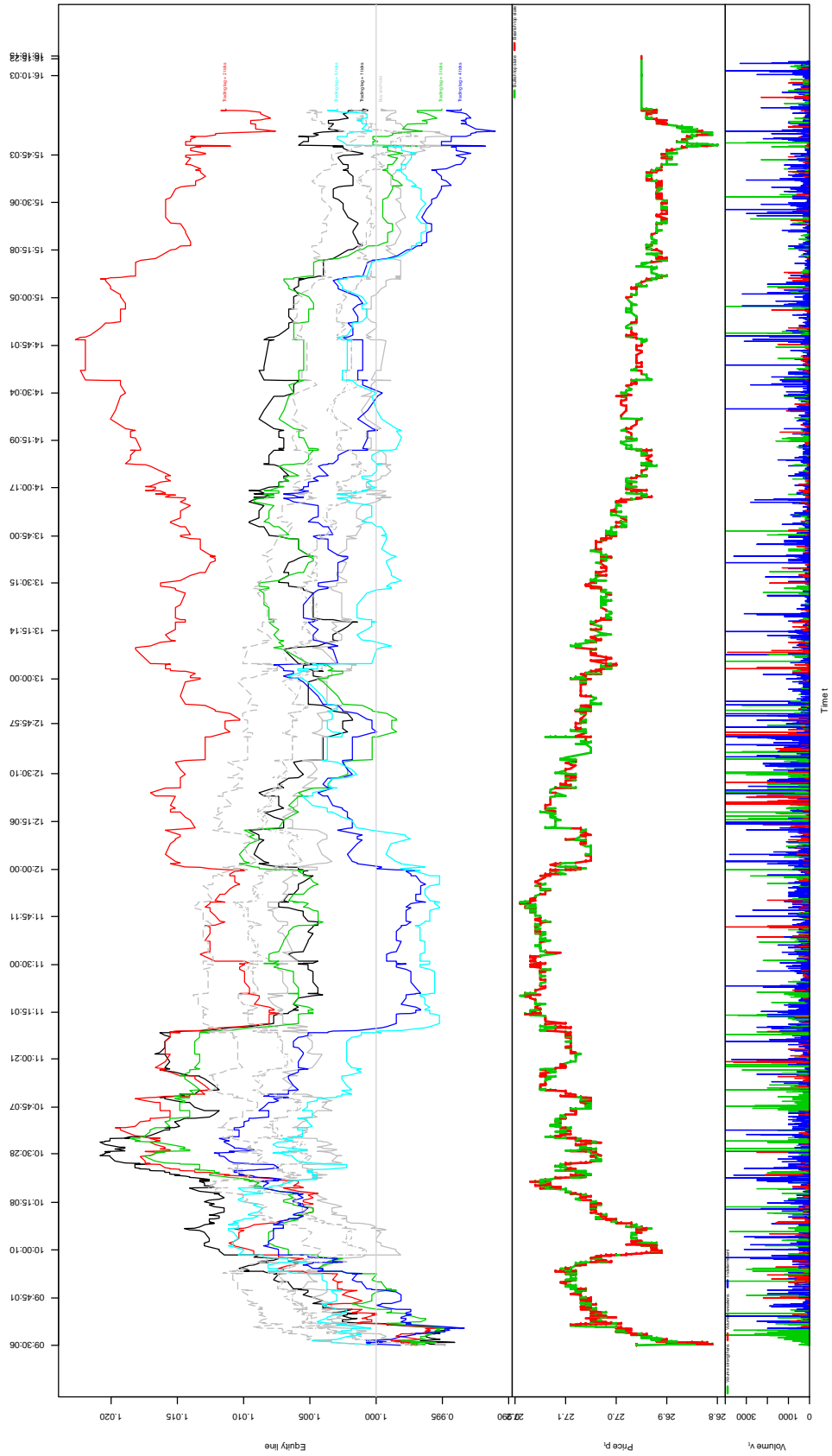
| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| Min | -4.51 | -21.54 | -42.41 | -25.71 | -7.47 | -5.76 | -6.09 |
| Mean | -0.01 | -0.18 | -0.95 | -0.17 | 0.30 | 0.44 | 0.45 |
| Median | -0.02 | 0.07 | -0.29 | -0.00 | 0.22 | 0.35 | 0.46 |
| Max | 5.82 | 20.56 | 12.90 | 11.27 | 8.08 | 8.23 | 5.71 |
| SD | 1.69 | 4.11 | 4.71 | 3.25 | 2.08 | 1.95 | 1.89 |
| IQR | 1.96 | 3.70 | 3.11 | 3.12 | 2.56 | 2.76 | 2.25 |

Table 6: Summary statistics of daily return originated in the HHMM trading strategy for different levels of lags. Returns expressed as percentages.

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| Buy-Hold | 1.00 | 0.07 | 0.00 | -0.02 | 0.06 | -0.07 | 0.04 |
| HHMM (0 lag) | 0.07 | 1.00 | 0.76 | 0.62 | 0.50 | 0.25 | -0.07 |
| HHMM (1 lag) | 0.00 | 0.76 | 1.00 | 0.82 | 0.55 | 0.27 | -0.08 |
| HHMM (2 lag) | -0.02 | 0.62 | 0.82 | 1.00 | 0.71 | 0.40 | 0.07 |
| HHMM (3 lag) | 0.06 | 0.50 | 0.55 | 0.71 | 1.00 | 0.62 | 0.30 |
| HHMM (4 lag) | -0.07 | 0.25 | 0.27 | 0.40 | 0.62 | 1.00 | 0.45 |
| HHMM (5 lag) | 0.04 | -0.07 | -0.08 | 0.07 | 0.30 | 0.45 | 1.00 |

Table 7: Correlation matrix of daily return originated in the HHMM trading strategy for different levels of lags.

## 3.7 Discussion

### 3.7.1 The statistical model

We find that the model proposed, while statistically simple, is highly expressive of financial domain knowledge. The hierarchical design creates a multi-level stochastic process that learns autocorrelations in different time scales. This accommodates for multi-resolution, a typical characteristic of financial datasets. We remark that, whereas HHMM offer a methodology to create a complex, highly non linear model for time series, estimation and inference can be achieved in reasonable complexity. The developments by Murphy and Paskin (2001) simplified cubic time into linear time inference, making inference feasible for high-frequency finance.

### 3.7.2 The financial application

We have special interest in the feature extraction procedure. They were cleverly designed to reproduce some of the most basic principles of technical analysis and, when applied to real data, they proved to be a powerful descriptor of price and volume movements. Nonetheless, we observe that the change in volume is the weakest component and provides with great opportunities for further enhancements. The contribution of the author should not be neglected: in general terms, the features are the most important factor in the success or the failure of a machine learning project (Domingos 2012).

---

# 4 References

Brock, William, Josef Lakonishok, and Blake LeBaron. 1992. "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns." *The Journal of Finance* 47 (5). Wiley Online Library: 1731–64.

Carpenter, Bob, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2016. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 20.

Chauvet, Marcelle, and Simon Potter. 2000. "Coincident and Leading Indicators of the Stock Market." *Journal of Empirical Finance* 7 (1). Elsevier: 87–111.

Domingos, Pedro. 2012. "A Few Useful Things to Know About Machine Learning." *Commun. ACM* 55 (10). New York, NY, USA: ACM: 78–87. doi:10.1145/2347736.2347755.

Fine, Shai, and Yoram Singer. 1998. "The Hierarchical Hidden Markov Model: Analysis and Applications."

Gallant, A Ronald, Peter E Rossi, and George Tauchen. 1992. "Stock Prices and Volume." *The Review of Financial Studies* 5 (2). Oxford University Press: 199–242.

Gelman, Andrew, and Donald B Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science.* JSTOR, 457–72.

Gordon, Stephen, and Pascal St-Amour. 2000. "A Preference Regime Model of Bull and Bear Markets." *The American Economic Review* 90 (4). JSTOR: 1019–33.

Karpoff, Jonathan M. 1987. "The Relation Between Price Changes and Trading Volume: A Survey." *Journal of Financial and Quantitative Analysis* 22 (1). Cambridge University Press: 109–26.

Lo, Andrew W, Harry Mamaysky, and Jiang Wang. 2000. "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation." *The Journal of Finance* 55 (4). Wiley

Online Library: 1705–65.

Lunde, Asger, and Allan G. Timmermann. 2004. "Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets."

Maheu, John M., and Thomas H. McCurdy. 2000. "Identifying Bull and Bear Markets in Stock Returns."

Murphy, Kevin P., and Mark A. Paskin. 2001. "Linear Time Inference in Hierarchical Hmms."

Ord, Tim. 2008. "The Secret Science of Price and Volume." *Master Traders: Strategies for Superior Returns from Today's Top Traders*. Wiley Online Library, 87–105.

Park, Cheol-Ho, and Scott H Irwin. 2007. "What Do We Know About the Profitability of Technical Analysis?" *Journal of Economic Surveys* 21 (4). Wiley Online Library: 786–826.

Sandoval, Javier, and Germán Hernández. 2015. "Computational Visual Analysis of the Order Book Dynamics for Creating High-Frequency Foreign Exchange Trading Strategies." *Procedia Computer Science* 51. Elsevier: 1593–1602.

Tayal, Aditya. 2009. "Regime Switching and Technical Trading with Dynamic Bayesian Networks in High-Frequency Stock Markets." Master's thesis, University of Waterloo.

Team, Stan Development. 2017. *Stan Modeling Language: User's Guide and Reference Manual: Version 2.15.0.*

Wisebourt, Shaul Sergey. 2011. "Hierarchical Hidden Markov Model of High-Frequency Market Regimes Using Trade Price and Limit Order Book Information." Master's thesis, University of Waterloo.

---

# 5 Appendix

## 5.1 Estimated parameters TSE:G 2007-05-04/2007-05-10.

|  | Mean | MCSE | SD | $q_{2.5\%}$ | $q_{10.0\%}$ | $q_{25.0\%}$ | $q_{50.0\%}$ | $q_{75.0\%}$ | $q_{90.0\%}$ | $q_{97.5\%}$ | ESS | $\hat{R}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\pi_1$ | 0.51 | 0.02 | 0.28 | 0.03 | 0.10 | 0.27 | 0.52 | 0.74 | 0.89 | 0.96 | 250.00 | 1.00 |
| $\pi_2$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $\pi_3$ | 0.49 | 0.02 | 0.28 | 0.04 | 0.11 | 0.26 | 0.48 | 0.73 | 0.90 | 0.97 | 250.00 | 1.00 |
| $\pi_4$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{11}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{12}$ | 0.46 | 0.01 | 0.11 | 0.19 | 0.31 | 0.39 | 0.47 | 0.53 | 0.58 | 0.64 | 119.70 | 1.00 |
| $a_{13}$ | 0.54 | 0.01 | 0.11 | 0.36 | 0.42 | 0.47 | 0.53 | 0.61 | 0.69 | 0.81 | 119.70 | 1.00 |
| $a_{14}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{21}$ | 1.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 250.00 | n.a. |
| $a_{22}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{23}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{24}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{31}$ | 0.09 | 0.00 | 0.06 | 0.00 | 0.02 | 0.04 | 0.08 | 0.13 | 0.17 | 0.24 | 198.87 | 1.00 |
| $a_{32}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{33}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{34}$ | 0.91 | 0.00 | 0.06 | 0.76 | 0.83 | 0.87 | 0.92 | 0.96 | 0.98 | 1.00 | 198.87 | 1.00 |
| $a_{41}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{42}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $a_{43}$ | 1.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 250.00 | n.a. |
| $a_{44}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | n.a. |
| $\phi_{11}$ | 0.01 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.02 | 0.02 | 99.42 | 1.01 |
| $\phi_{12}$ | 0.02 | 0.00 | 0.00 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 250.00 | 1.00 |
| $\phi_{13}$ | 0.01 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.03 | 102.05 | 1.02 |
| $\phi_{14}$ | 0.34 | 0.00 | 0.02 | 0.29 | 0.31 | 0.33 | 0.34 | 0.36 | 0.37 | 0.38 | 150.07 | 1.01 |
| $\phi_{15}$ | 0.22 | 0.00 | 0.03 | 0.17 | 0.19 | 0.21 | 0.23 | 0.24 | 0.26 | 0.27 | 171.58 | 1.01 |
| $\phi_{16}$ | 0.35 | 0.00 | 0.02 | 0.31 | 0.33 | 0.34 | 0.35 | 0.37 | 0.37 | 0.39 | 250.00 | 1.00 |
| $\phi_{17}$ | 0.03 | 0.00 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 0.04 | 250.00 | 1.00 |
| $\phi_{18}$ | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.02 | 144.00 | 1.00 |
| $\phi_{19}$ | 0.02 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 250.00 | 1.00 |
| $\phi_{21}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 66.44 | 1.01 |
| $\phi_{22}$ | 0.02 | 0.00 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 250.00 | 1.00 |
| $\phi_{23}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 250.00 | 1.00 |
| $\phi_{24}$ | 0.05 | 0.00 | 0.03 | 0.00 | 0.01 | 0.02 | 0.05 | 0.07 | 0.09 | 0.12 | 102.30 | 1.00 |
| $\phi_{25}$ | 0.80 | 0.00 | 0.04 | 0.72 | 0.75 | 0.78 | 0.80 | 0.83 | 0.85 | 0.87 | 90.82 | 1.00 |
| $\phi_{26}$ | 0.02 | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.02 | 0.03 | 0.05 | 0.07 | 194.25 | 1.00 |
| $\phi_{27}$ | 0.08 | 0.00 | 0.01 | 0.06 | 0.07 | 0.07 | 0.08 | 0.09 | 0.09 | 0.10 | 250.00 | 1.00 |
| $\phi_{28}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | 1.00 |
| $\phi_{29}$ | 0.02 | 0.00 | 0.00 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 250.00 | 1.00 |
| $\phi_{31}$ | 0.01 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 52.96 | 1.00 |
| $\phi_{32}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 250.00 | 1.00 |
| $\phi_{33}$ | 0.03 | 0.00 | 0.01 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 0.04 | 250.00 | 1.00 |
| $\phi_{34}$ | 0.36 | 0.00 | 0.03 | 0.30 | 0.32 | 0.33 | 0.35 | 0.38 | 0.39 | 0.41 | 111.49 | 1.00 |
| $\phi_{35}$ | 0.20 | 0.00 | 0.04 | 0.14 | 0.15 | 0.17 | 0.20 | 0.22 | 0.24 | 0.28 | 99.99 | 1.00 |
| $\phi_{36}$ | 0.39 | 0.00 | 0.02 | 0.35 | 0.36 | 0.38 | 0.39 | 0.40 | 0.42 | 0.43 | 227.14 | 1.00 |
| $\phi_{37}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 250.00 | 1.00 |
| $\phi_{38}$ | 0.02 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 250.00 | 1.00 |
| $\phi_{39}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 250.00 | 1.00 |
| $\phi_{41}$ | 0.02 | 0.00 | 0.01 | 0.00 | 0.01 | 0.01 | 0.02 | 0.02 | 0.03 | 0.03 | 75.13 | 1.00 |
| $\phi_{42}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 250.00 | 1.00 |
| $\phi_{43}$ | 0.06 | 0.00 | 0.01 | 0.03 | 0.04 | 0.05 | 0.06 | 0.06 | 0.07 | 0.07 | 124.34 | 1.02 |
| $\phi_{44}$ | 0.02 | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.02 | 0.03 | 0.05 | 0.06 | 178.86 | 1.01 |
| $\phi_{45}$ | 0.88 | 0.00 | 0.02 | 0.83 | 0.85 | 0.86 | 0.88 | 0.90 | 0.91 | 0.93 | 110.86 | 1.02 |
| $\phi_{46}$ | 0.01 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 151.90 | 1.00 |
| $\phi_{47}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 250.00 | 1.00 |
| $\phi_{48}$ | 0.01 | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 110.72 | 1.00 |
| $\phi_{49}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 250.00 | 1.00 |

Table 8: Statistics summary for the distributions of estimated parameters for TSE:G 2007-05-04 09:30:00/2007-05-10 16:30:00. MCSE means Monte Carlo Standard Error, SD means (posteriori) Standard Deviation and ESS means Effective Sample Size.

## 5.2 Out of sample trading performance for some selected assets

### 5.2.1 BBDb.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | -0.01 | -0.06 | -0.05 | -0.02 | -0.00 | -0.02 | -0.04 |
| 2007-05-09 | -0.01 | 0.00 | -0.06 | -0.01 | -0.04 | 0.03 | 0.00 |
| 2007-05-10 | -0.01 | -0.01 | -0.02 | 0.02 | 0.03 | 0.02 | 0.01 |
| 2007-05-11 | 0.02 | -0.05 | -0.05 | -0.03 | 0.01 | -0.01 | -0.00 |
| 2007-05-14 | -0.00 | -0.12 | -0.14 | -0.03 | -0.03 | -0.01 | -0.01 |
| 2007-05-15 | 0.00 | -0.03 | -0.00 | 0.03 | -0.00 | -0.02 | 0.02 |
| 2007-05-16 | 0.01 | -0.04 | -0.15 | -0.07 | 0.01 | 0.01 | 0.03 |
| 2007-05-17 | 0.01 | 0.00 | -0.03 | -0.01 | 0.01 | 0.02 | 0.01 |
| 2007-05-18 | -0.01 | -0.06 | -0.12 | -0.04 | 0.00 | 0.04 | 0.02 |
| 2007-05-22 | -0.01 | -0.05 | -0.09 | -0.06 | -0.05 | -0.01 | 0.01 |
| 2007-05-23 | -0.02 | -0.10 | -0.18 | -0.04 | 0.01 | 0.03 | 0.05 |
| 2007-05-24 | 0.00 | -0.13 | -0.08 | 0.03 | -0.00 | 0.03 | -0.03 |
| 2007-05-25 | 0.01 | -0.16 | -0.10 | -0.06 | -0.04 | -0.06 | 0.05 |
| 2007-05-28 | 0.01 | -0.10 | -0.09 | -0.04 | -0.02 | -0.02 | 0.04 |
| 2007-05-29 | 0.04 | -0.22 | -0.42 | -0.26 | -0.07 | -0.04 | 0.05 |
| 2007-05-30 | 0.01 | -0.11 | -0.21 | -0.11 | -0.04 | 0.02 | 0.04 |
| 2007-05-31 | -0.03 | 0.01 | 0.00 | 0.09 | 0.05 | 0.01 | 0.03 |
| Total | -0.01 | -0.73 | -0.87 | -0.49 | -0.17 | 0.03 | 0.34 |
| Min | -0.03 | -0.22 | -0.42 | -0.26 | -0.07 | -0.06 | -0.04 |
| Mean | -0.00 | -0.07 | -0.11 | -0.04 | -0.01 | 0.00 | 0.02 |
| Median | 0.00 | -0.06 | -0.09 | -0.03 | -0.00 | 0.01 | 0.02 |
| Max | 0.04 | 0.01 | 0.00 | 0.09 | 0.05 | 0.04 | 0.05 |
| SD | 0.02 | 0.06 | 0.10 | 0.07 | 0.03 | 0.03 | 0.03 |

Table 9: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). BBDb.TO



Figure 10: Out of sample equity line for BBDb.TO [2007-05-31 09:30:00/2007-05-31 16:30:00]

### 5.2.2 BCE.TO

|  | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | 0.01 | -0.02 | 0.02 | 0.00 | -0.02 | 0.01 | 0.01 |
| 2007-05-09 | -0.01 | -0.04 | -0.01 | -0.02 | 0.01 | 0.01 | 0.01 |
| 2007-05-10 | 0.00 | -0.02 | -0.01 | 0.02 | 0.01 | 0.00 | 0.01 |
| 2007-05-11 | -0.00 | 0.00 | -0.01 | 0.01 | -0.01 | -0.01 | -0.01 |
| 2007-05-14 | -0.00 | -0.02 | -0.02 | -0.02 | -0.01 | 0.00 | -0.01 |
| 2007-05-15 | 0.00 | -0.02 | -0.01 | 0.01 | -0.01 | 0.01 | 0.02 |
| 2007-05-16 | 0.00 | -0.04 | -0.03 | -0.03 | -0.02 | 0.00 | 0.01 |
| 2007-05-17 | -0.00 | -0.02 | -0.01 | 0.02 | -0.01 | 0.01 | 0.02 |
| 2007-05-18 | 0.02 | -0.03 | -0.02 | -0.02 | -0.02 | -0.01 | -0.02 |
| 2007-05-22 | 0.03 | -0.03 | 0.02 | -0.01 | -0.00 | -0.01 | 0.01 |
| 2007-05-23 | 0.02 | -0.01 | -0.04 | 0.00 | -0.03 | -0.02 | -0.02 |
| 2007-05-24 | 0.00 | -0.05 | -0.00 | -0.01 | 0.00 | -0.00 | 0.00 |
| 2007-05-25 | 0.00 | -0.02 | 0.02 | 0.02 | 0.00 | -0.00 | -0.00 |
| 2007-05-28 | 0.00 | -0.00 | -0.00 | -0.01 | -0.01 | -0.00 | -0.00 |
| 2007-05-29 | -0.00 | -0.02 | -0.02 | -0.02 | -0.01 | -0.00 | -0.01 |
| 2007-05-30 | 0.01 | -0.01 | 0.00 | -0.02 | -0.01 | -0.02 | -0.00 |
| 2007-05-31 | -0.00 | -0.01 | 0.02 | 0.02 | 0.03 | 0.01 | 0.00 |
| Total | 0.08 | -0.31 | -0.12 | -0.07 | -0.09 | -0.02 | 0.02 |
| Min | -0.01 | -0.05 | -0.04 | -0.03 | -0.03 | -0.02 | -0.02 |
| Mean | 0.00 | -0.02 | -0.01 | -0.00 | -0.01 | -0.00 | 0.00 |
| Median | 0.00 | -0.02 | -0.01 | -0.01 | -0.01 | -0.00 | 0.00 |
| Max | 0.03 | 0.00 | 0.02 | 0.02 | 0.03 | 0.01 | 0.02 |
| SD | 0.01 | 0.01 | 0.02 | 0.02 | 0.01 | 0.01 | 0.01 |

Table 10: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). BCE.TO
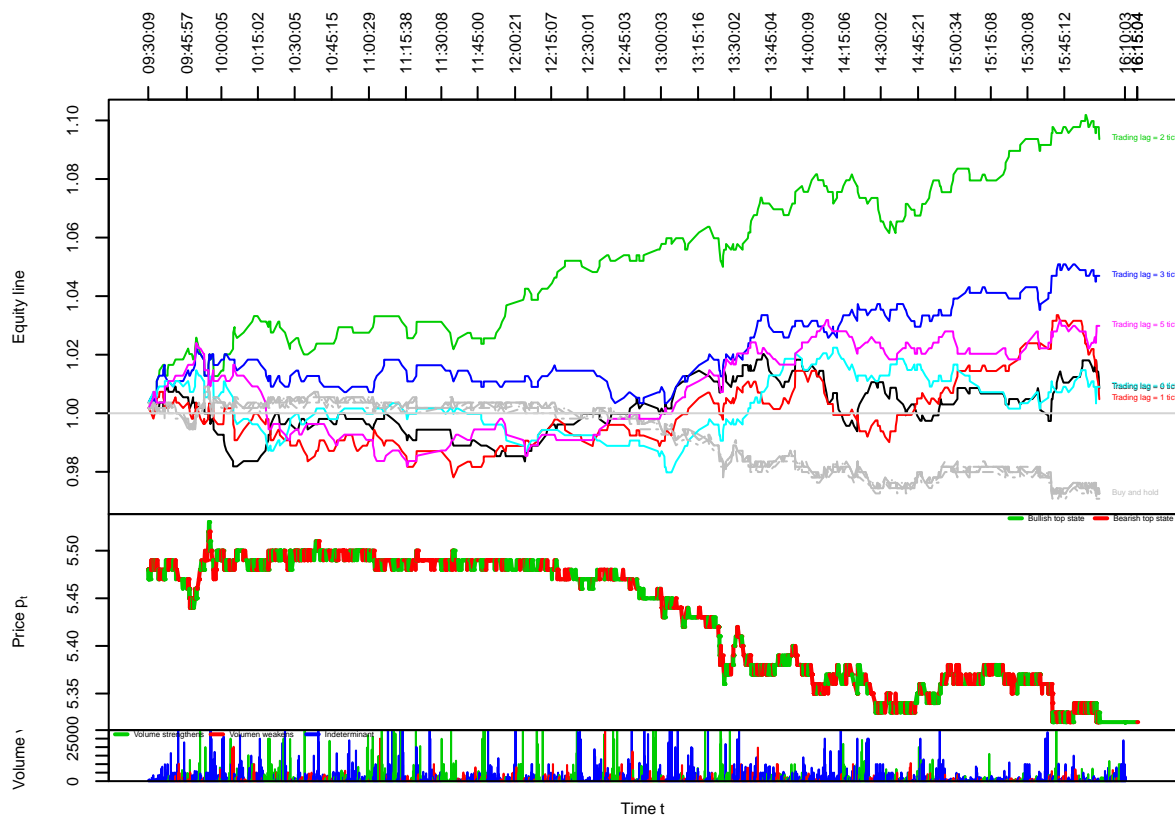


Figure 11: Out of sample equity line for BCE.TO [2007-05-31 09:30:00/2007-05-31 16:30:00]

### 5.2.3 CTCa.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | -0.01 | -0.00 | -0.00 | 0.00 | -0.01 | 0.00 | -0.00 |
| 2007-05-09 | -0.01 | -0.02 | -0.02 | -0.00 | 0.00 | 0.00 | -0.01 |
| 2007-05-10 | 0.05 | 0.03 | 0.02 | -0.01 | 0.03 | 0.04 | 0.06 |
| 2007-05-11 | 0.02 | 0.04 | 0.03 | 0.04 | 0.03 | 0.04 | 0.03 |
| 2007-05-14 | -0.00 | 0.04 | 0.01 | 0.00 | 0.02 | 0.00 | -0.01 |
| 2007-05-15 | 0.00 | -0.01 | -0.01 | 0.02 | 0.01 | 0.01 | -0.01 |
| 2007-05-16 | 0.02 | -0.02 | 0.01 | 0.02 | 0.02 | 0.01 | 0.04 |
| 2007-05-17 | 0.01 | 0.01 | 0.03 | 0.04 | 0.03 | 0.02 | 0.03 |
| 2007-05-18 | -0.01 | 0.02 | 0.02 | -0.00 | 0.00 | -0.01 | 0.00 |
| 2007-05-22 | 0.00 | 0.01 | -0.00 | -0.01 | 0.00 | 0.02 | 0.01 |
| 2007-05-23 | 0.01 | -0.01 | -0.01 | -0.00 | 0.01 | 0.02 | 0.00 |
| 2007-05-24 | -0.01 | -0.01 | -0.00 | 0.01 | 0.01 | 0.01 | 0.01 |
| 2007-05-25 | 0.01 | -0.00 | 0.02 | 0.01 | 0.00 | 0.01 | -0.01 |
| 2007-05-28 | 0.00 | 0.02 | -0.03 | 0.00 | 0.01 | 0.01 | 0.02 |
| 2007-05-29 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.00 | 0.02 |
| 2007-05-30 | 0.00 | 0.03 | 0.01 | -0.01 | -0.01 | -0.04 | -0.01 |
| 2007-05-31 | -0.00 | -0.01 | -0.02 | -0.02 | -0.03 | -0.00 | 0.01 |
| | | | | | | | |
| Total | 0.06 | 0.13 | 0.03 | 0.07 | 0.13 | 0.15 | 0.21 |
| Min | -0.01 | -0.02 | -0.03 | -0.02 | -0.03 | -0.04 | -0.01 |
| Mean | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 |
| Median | 0.00 | -0.00 | -0.00 | 0.00 | 0.01 | 0.01 | 0.01 |
| Max | 0.05 | 0.04 | 0.03 | 0.04 | 0.03 | 0.04 | 0.06 |
| SD | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |

Table 11: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). CTCa.TO
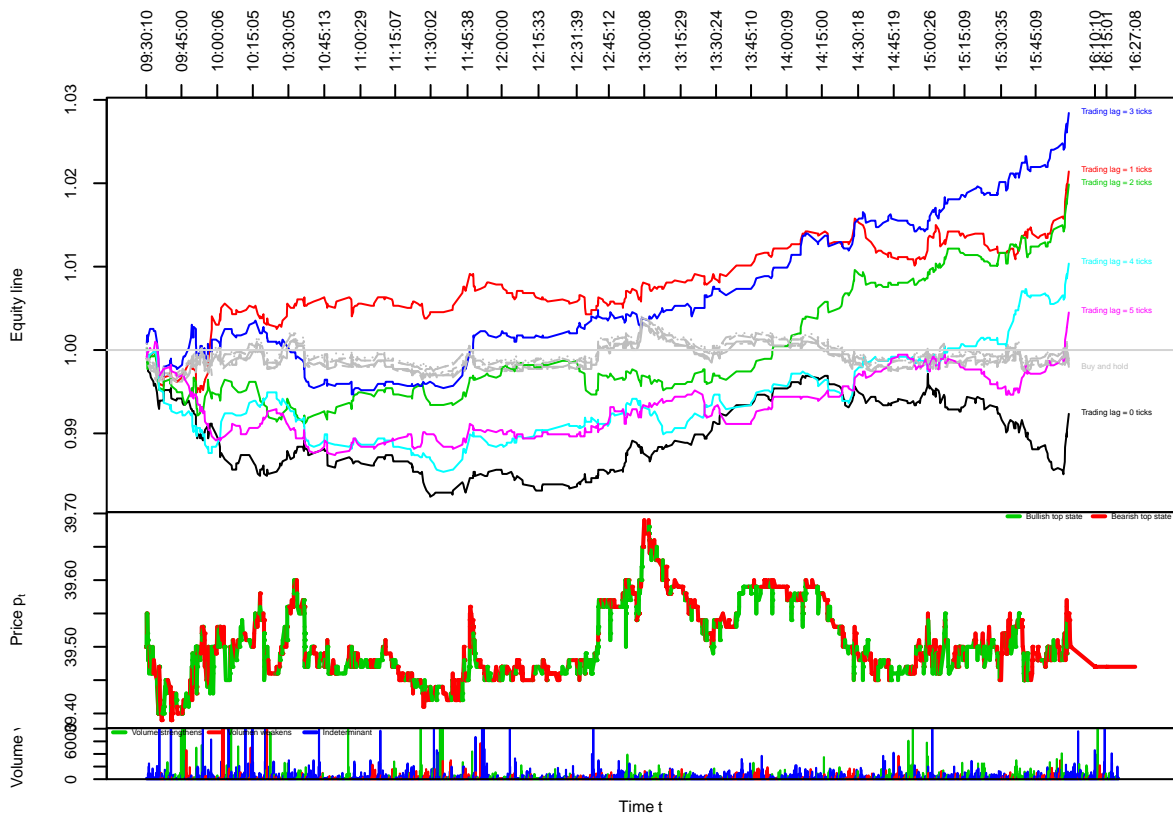


Figure 12: Out of sample equity line for CTCa.TO [2007-05-11 09:30:00/2007-05-11 16:30:00]

### 5.2.4   ECA.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | 0.01 | -0.02 | -0.02 | -0.02 | -0.01 | -0.01 | 0.01 |
| 2007-05-09 | -0.01 | -0.00 | -0.00 | 0.01 | 0.00 | -0.02 | 0.01 |
| 2007-05-10 | 0.00 | -0.01 | -0.04 | -0.04 | -0.03 | -0.04 | -0.03 |
| 2007-05-11 | 0.05 | 0.00 | -0.02 | -0.02 | 0.01 | -0.03 | -0.02 |
| 2007-05-14 | -0.01 | -0.02 | -0.01 | -0.00 | 0.00 | 0.02 | -0.02 |
| 2007-05-15 | 0.00 | 0.03 | -0.03 | -0.01 | -0.02 | -0.03 | -0.02 |
| 2007-05-16 | 0.02 | 0.02 | 0.03 | 0.03 | 0.02 | 0.02 | 0.01 |
| 2007-05-17 | 0.02 | 0.03 | 0.04 | 0.02 | 0.02 | 0.02 | 0.02 |
| 2007-05-18 | 0.00 | 0.02 | 0.02 | 0.02 | 0.01 | -0.00 | 0.03 |
| 2007-05-22 | -0.01 | -0.00 | 0.00 | -0.00 | -0.00 | -0.01 | -0.01 |
| 2007-05-23 | -0.01 | 0.01 | 0.01 | 0.00 | 0.01 | 0.02 | 0.02 |
| 2007-05-24 | -0.02 | -0.04 | -0.00 | 0.01 | 0.02 | 0.02 | 0.03 |
| 2007-05-25 | 0.00 | 0.04 | 0.05 | 0.03 | 0.03 | 0.02 | 0.02 |
| 2007-05-28 | -0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.02 | -0.00 |
| 2007-05-29 | -0.01 | 0.04 | 0.01 | -0.00 | -0.01 | -0.02 | -0.03 |
| 2007-05-30 | 0.03 | 0.02 | 0.02 | -0.00 | 0.02 | 0.01 | 0.01 |
| 2007-05-31 | -0.01 | 0.06 | 0.02 | 0.02 | 0.03 | 0.03 | 0.01 |
| Total | 0.06 | 0.24 | 0.09 | 0.05 | 0.12 | 0.03 | 0.04 |
| Min | -0.02 | -0.04 | -0.04 | -0.04 | -0.03 | -0.04 | -0.03 |
| Mean | 0.00 | 0.01 | 0.01 | 0.00 | 0.01 | 0.00 | 0.00 |
| Median | 0.00 | 0.02 | 0.01 | 0.00 | 0.01 | 0.01 | 0.01 |
| Max | 0.05 | 0.06 | 0.05 | 0.03 | 0.03 | 0.03 | 0.03 |
| SD | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |

Table 12: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). ECA.TO
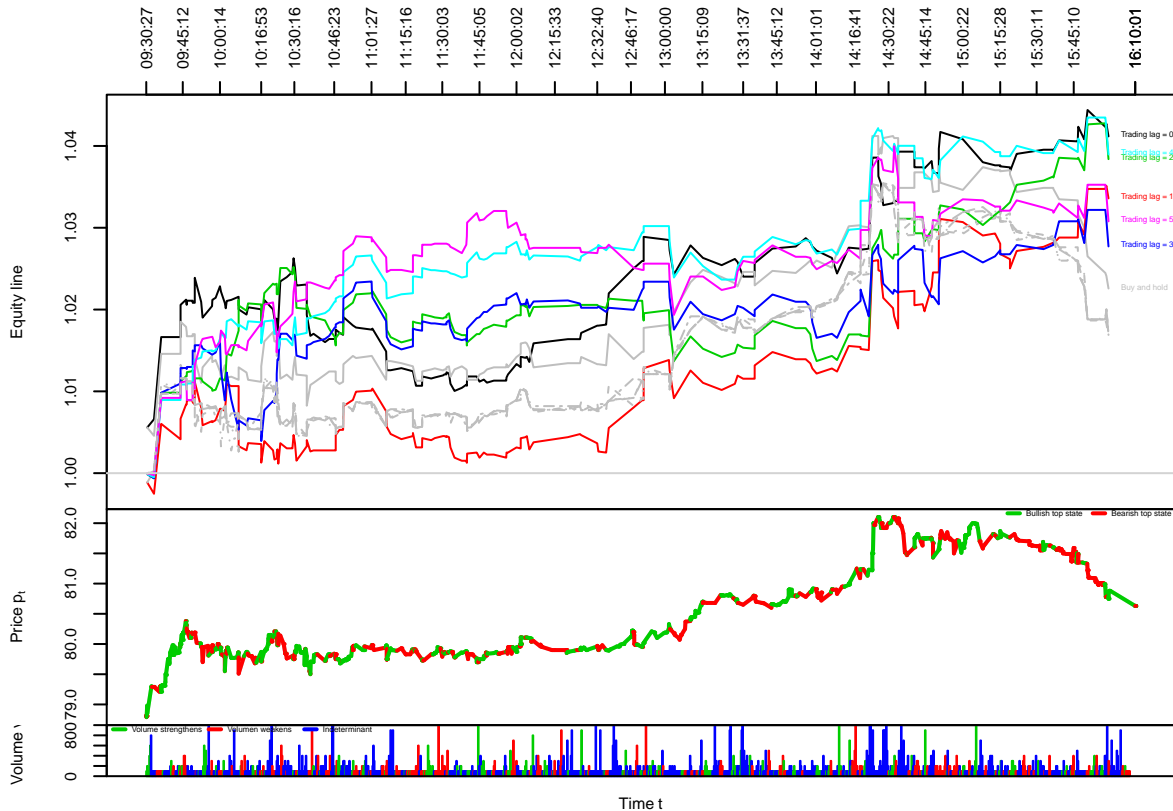


Figure 13: Out of sample equity line for ECA.TO [2007-05-25 09:30:00/2007-05-25 16:30:00]

### 5.2.5 G.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | -0.01 | 0.04 | -0.02 | -0.01 | 0.01 | 0.01 | 0.02 |
| 2007-05-09 | -0.00 | 0.04 | 0.00 | 0.02 | 0.02 | 0.01 | 0.01 |
| 2007-05-10 | -0.00 | 0.04 | 0.01 | -0.01 | 0.02 | 0.02 | 0.00 |
| 2007-05-11 | -0.00 | 0.00 | 0.00 | 0.01 | -0.00 | -0.01 | 0.00 |
| 2007-05-14 | -0.03 | 0.03 | -0.01 | 0.01 | -0.01 | -0.00 | -0.01 |
| 2007-05-15 | -0.00 | 0.03 | -0.00 | 0.00 | 0.02 | 0.02 | 0.00 |
| 2007-05-16 | -0.00 | 0.05 | -0.01 | -0.00 | 0.01 | 0.03 | -0.02 |
| 2007-05-17 | -0.00 | -0.02 | 0.00 | 0.02 | 0.00 | -0.03 | -0.00 |
| 2007-05-18 | 0.01 | -0.01 | 0.01 | 0.00 | 0.02 | 0.01 | 0.02 |
| 2007-05-22 | -0.02 | -0.02 | -0.02 | 0.02 | 0.01 | 0.02 | 0.02 |
| 2007-05-23 | -0.01 | 0.02 | -0.00 | -0.01 | -0.01 | 0.01 | 0.01 |
| 2007-05-24 | -0.03 | 0.02 | -0.00 | -0.01 | -0.01 | -0.00 | -0.02 |
| 2007-05-25 | 0.00 | -0.01 | -0.01 | 0.00 | 0.01 | -0.02 | -0.02 |
| 2007-05-28 | -0.01 | -0.02 | -0.01 | 0.01 | -0.01 | -0.01 | 0.02 |
| 2007-05-29 | -0.02 | -0.02 | -0.01 | -0.02 | -0.00 | 0.02 | 0.01 |
| 2007-05-30 | 0.01 | -0.02 | -0.03 | -0.03 | -0.01 | -0.04 | -0.02 |
| 2007-05-31 | 0.04 | 0.00 | -0.01 | -0.01 | -0.02 | -0.01 | -0.03 |
| | | | | | | | |
| Total | -0.09 | 0.17 | -0.12 | -0.01 | 0.04 | 0.02 | -0.02 |
| Min | -0.03 | -0.02 | -0.03 | -0.03 | -0.02 | -0.04 | -0.03 |
| Mean | -0.01 | 0.01 | -0.01 | -0.00 | 0.00 | 0.00 | -0.00 |
| Median | -0.00 | 0.00 | -0.01 | 0.00 | 0.00 | 0.01 | 0.00 |
| Max | 0.04 | 0.05 | 0.01 | 0.02 | 0.02 | 0.03 | 0.02 |
| SD | 0.02 | 0.03 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 |

Table 13: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). G.TO
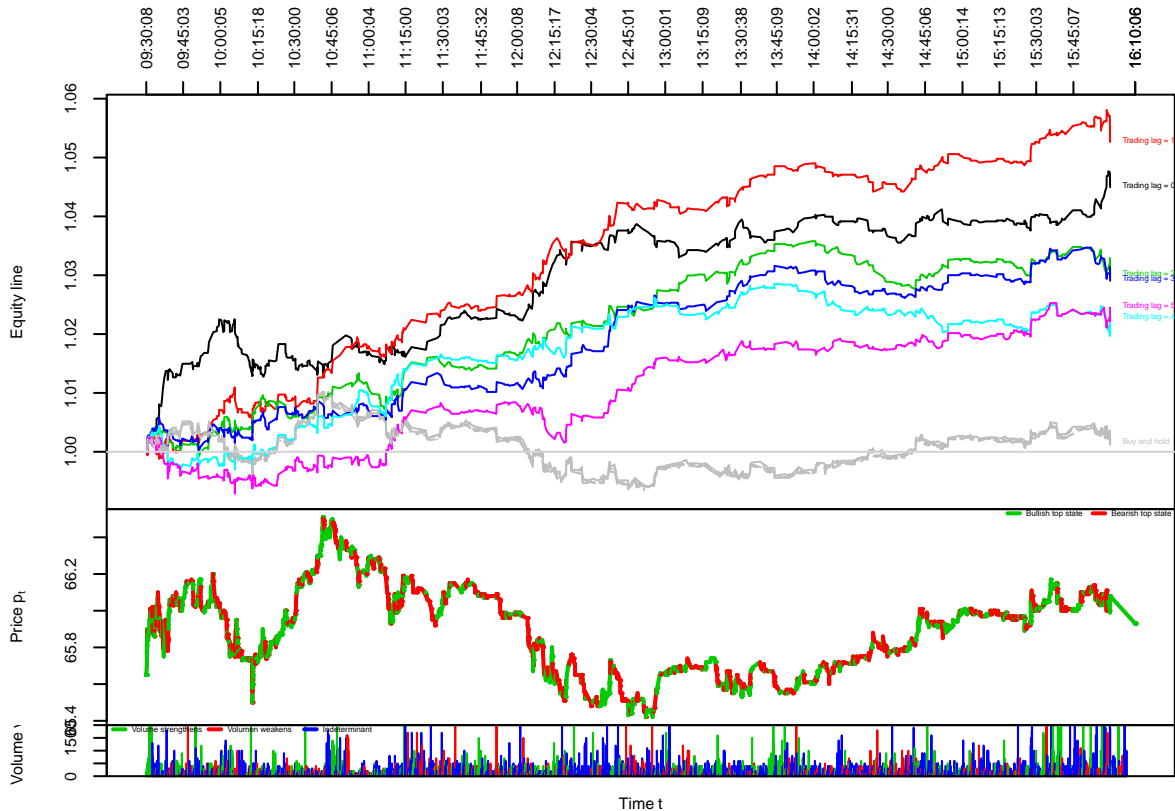


Figure 14: Out of sample equity line for G.TO [2007-05-10 09:30:00/2007-05-10 16:30:00]

### 5.2.6   K.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | -0.02 | -0.00 | -0.01 | 0.03 | 0.02 | 0.01 | 0.02 |
| 2007-05-09 | -0.01 | -0.05 | -0.04 | 0.01 | 0.02 | 0.01 | -0.00 |
| 2007-05-10 | -0.01 | 0.00 | 0.01 | 0.04 | 0.02 | 0.00 | -0.01 |
| 2007-05-11 | 0.01 | -0.02 | -0.05 | -0.03 | -0.00 | -0.02 | -0.01 |
| 2007-05-14 | -0.02 | 0.02 | 0.00 | 0.03 | 0.03 | 0.04 | -0.03 |
| 2007-05-15 | -0.01 | 0.05 | 0.01 | 0.04 | 0.04 | 0.03 | 0.02 |
| 2007-05-16 | -0.01 | 0.02 | -0.05 | -0.04 | -0.03 | -0.03 | -0.01 |
| 2007-05-17 | -0.00 | 0.01 | 0.01 | 0.03 | 0.03 | 0.03 | 0.01 |
| 2007-05-18 | 0.00 | -0.02 | -0.02 | -0.03 | 0.01 | 0.01 | 0.00 |
| 2007-05-22 | -0.02 | 0.01 | 0.00 | 0.01 | 0.01 | 0.02 | 0.01 |
| 2007-05-23 | 0.01 | -0.00 | -0.01 | -0.02 | 0.01 | 0.02 | -0.00 |
| 2007-05-24 | -0.03 | -0.10 | -0.08 | -0.07 | -0.02 | 0.01 | 0.01 |
| 2007-05-25 | -0.01 | -0.01 | 0.00 | -0.04 | -0.02 | -0.00 | 0.00 |
| 2007-05-28 | 0.00 | -0.00 | -0.01 | -0.00 | 0.01 | -0.01 | 0.01 |
| 2007-05-29 | -0.03 | -0.03 | 0.02 | 0.01 | 0.04 | 0.02 | 0.03 |
| 2007-05-30 | -0.00 | -0.00 | -0.03 | -0.02 | -0.01 | -0.01 | -0.01 |
| 2007-05-31 | 0.04 | 0.03 | -0.01 | -0.01 | -0.01 | -0.02 | -0.00 |
| | | | | | | | |
| Total | -0.10 | -0.11 | -0.22 | -0.07 | 0.14 | 0.12 | 0.03 |
| Min | -0.03 | -0.10 | -0.08 | -0.07 | -0.03 | -0.03 | -0.03 |
| Mean | -0.01 | -0.01 | -0.01 | -0.00 | 0.01 | 0.01 | 0.00 |
| Median | -0.01 | -0.00 | -0.01 | -0.00 | 0.01 | 0.01 | 0.00 |
| Max | 0.04 | 0.05 | 0.02 | 0.04 | 0.04 | 0.04 | 0.03 |
| SD | 0.02 | 0.03 | 0.03 | 0.03 | 0.02 | 0.02 | 0.01 |

Table 14: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). K.TO
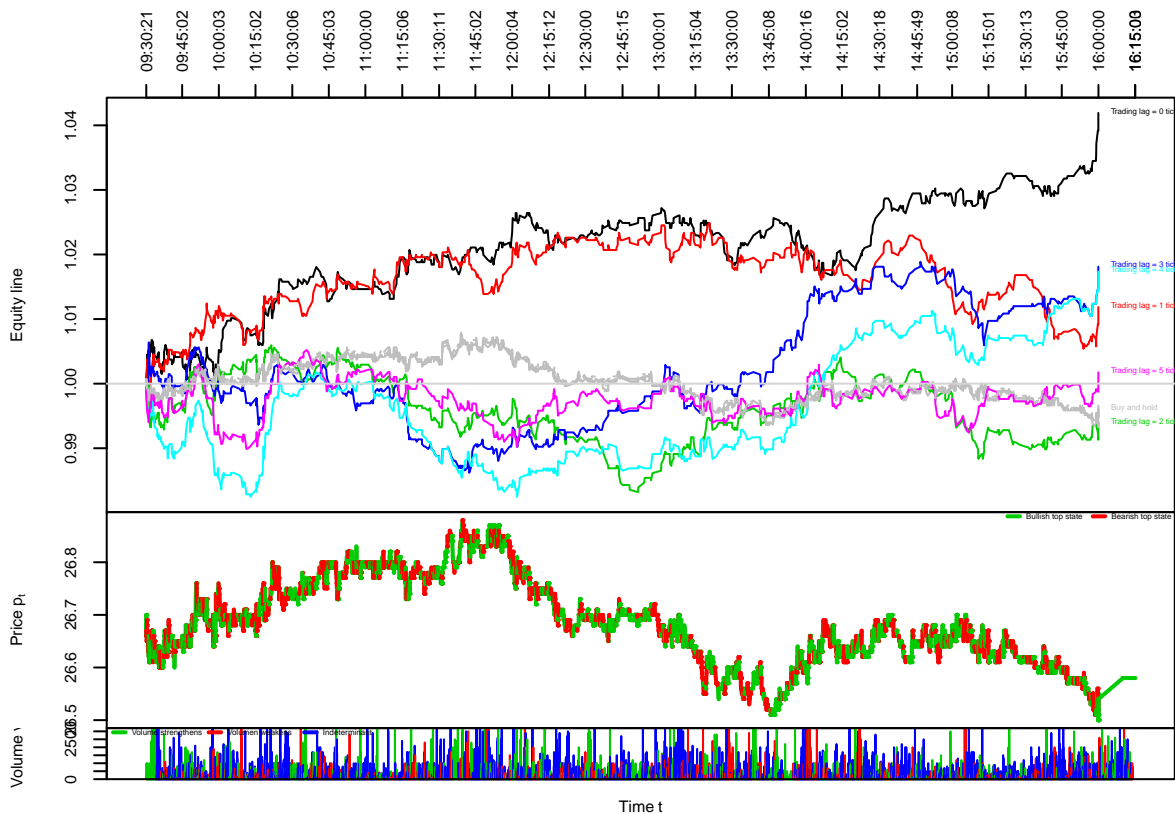


Figure 15: Out of sample equity line for K.TO [2007-05-29 09:30:00/2007-05-29 16:30:00]

### 5.2.7 MGa.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | -0.02 | -0.05 | 0.00 | 0.02 | 0.04 | 0.01 | -0.01 |
| 2007-05-09 | 0.00 | -0.01 | -0.02 | -0.02 | 0.01 | 0.01 | 0.01 |
| 2007-05-10 | 0.06 | 0.21 | 0.13 | 0.11 | 0.08 | 0.01 | 0.01 |
| 2007-05-11 | 0.01 | 0.01 | 0.00 | -0.00 | -0.01 | -0.01 | 0.01 |
| 2007-05-14 | -0.03 | -0.03 | 0.01 | 0.01 | 0.02 | 0.03 | 0.01 |
| 2007-05-15 | -0.02 | 0.02 | -0.01 | 0.01 | 0.02 | -0.02 | -0.02 |
| 2007-05-16 | 0.01 | 0.02 | 0.00 | 0.01 | 0.01 | -0.01 | 0.00 |
| 2007-05-17 | -0.01 | 0.08 | 0.07 | 0.08 | 0.05 | 0.00 | 0.00 |
| 2007-05-18 | 0.01 | 0.03 | 0.04 | 0.03 | 0.02 | 0.03 | 0.02 |
| 2007-05-22 | 0.02 | 0.03 | 0.03 | 0.02 | 0.03 | 0.02 | 0.03 |
| 2007-05-23 | 0.01 | 0.05 | 0.03 | 0.01 | -0.01 | -0.02 | -0.00 |
| 2007-05-24 | 0.02 | -0.03 | 0.01 | 0.03 | 0.03 | 0.03 | 0.02 |
| 2007-05-25 | -0.00 | 0.00 | 0.01 | 0.02 | 0.01 | 0.00 | -0.01 |
| 2007-05-28 | 0.01 | 0.00 | 0.02 | 0.00 | -0.02 | -0.01 | -0.01 |
| 2007-05-29 | -0.01 | 0.01 | 0.00 | 0.01 | 0.00 | 0.01 | -0.02 |
| 2007-05-30 | 0.00 | 0.01 | 0.00 | -0.02 | -0.01 | 0.01 | -0.00 |
| 2007-05-31 | 0.01 | 0.06 | 0.04 | 0.03 | 0.03 | 0.03 | 0.02 |
| Total | 0.07 | 0.48 | 0.45 | 0.38 | 0.32 | 0.11 | 0.07 |
| Min | -0.03 | -0.05 | -0.02 | -0.02 | -0.02 | -0.02 | -0.02 |
| Mean | 0.00 | 0.02 | 0.02 | 0.02 | 0.02 | 0.01 | 0.00 |
| Median | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 | 0.00 |
| Max | 0.06 | 0.21 | 0.13 | 0.11 | 0.08 | 0.03 | 0.03 |
| SD | 0.02 | 0.06 | 0.03 | 0.03 | 0.03 | 0.02 | 0.02 |

Table 15: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). MGa.TO
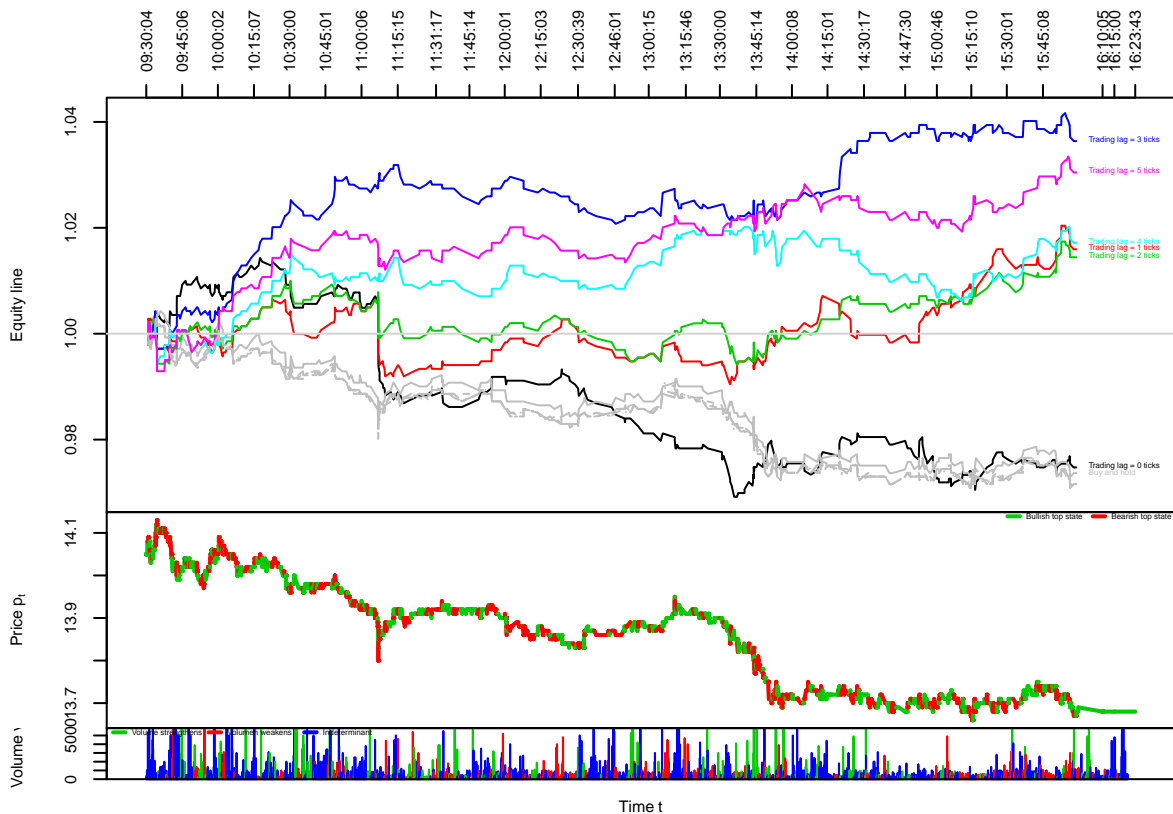


Figure 16: Out of sample equity line for MGa.TO [2007-05-10 09:30:00/2007-05-10 16:30:00]

### 5.2.8 NXY.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | 0.00 | 0.05 | -0.01 | 0.02 | 0.01 | 0.00 | 0.01 |
| 2007-05-09 | -0.01 | 0.04 | 0.03 | 0.03 | 0.04 | 0.04 | 0.04 |
| 2007-05-10 | -0.01 | -0.07 | -0.04 | -0.01 | 0.02 | -0.01 | 0.03 |
| 2007-05-11 | 0.02 | -0.00 | -0.04 | -0.02 | -0.00 | -0.00 | 0.01 |
| 2007-05-14 | -0.02 | -0.04 | -0.00 | 0.01 | 0.02 | 0.02 | 0.04 |
| 2007-05-15 | -0.01 | -0.01 | -0.02 | -0.01 | -0.00 | -0.00 | -0.01 |
| 2007-05-16 | 0.02 | -0.01 | -0.00 | -0.03 | 0.03 | 0.06 | 0.03 |
| 2007-05-17 | 0.02 | 0.03 | 0.00 | -0.00 | -0.01 | -0.02 | -0.02 |
| 2007-05-18 | -0.01 | -0.04 | -0.06 | -0.06 | -0.03 | -0.00 | 0.01 |
| 2007-05-22 | -0.02 | 0.01 | 0.01 | -0.01 | 0.01 | 0.01 | 0.01 |
| 2007-05-23 | -0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.04 | 0.02 |
| 2007-05-24 | -0.03 | 0.03 | 0.02 | 0.00 | 0.02 | 0.03 | 0.01 |
| 2007-05-25 | -0.02 | -0.01 | -0.05 | -0.03 | -0.01 | -0.01 | -0.01 |
| 2007-05-28 | 0.00 | 0.01 | 0.01 | 0.02 | 0.03 | 0.02 | 0.02 |
| 2007-05-29 | -0.00 | -0.00 | -0.01 | -0.01 | 0.01 | -0.01 | -0.02 |
| 2007-05-30 | 0.02 | -0.01 | -0.00 | 0.02 | 0.01 | 0.01 | -0.03 |
| 2007-05-31 | -0.01 | 0.02 | -0.00 | -0.00 | 0.01 | -0.00 | 0.00 |
| Total | -0.07 | 0.01 | -0.16 | -0.08 | 0.18 | 0.17 | 0.14 |
| Min | -0.03 | -0.07 | -0.06 | -0.06 | -0.03 | -0.02 | -0.03 |
| Mean | -0.00 | 0.00 | -0.01 | -0.00 | 0.01 | 0.01 | 0.01 |
| Median | -0.01 | -0.00 | -0.00 | -0.00 | 0.01 | 0.00 | 0.01 |
| Max | 0.02 | 0.05 | 0.03 | 0.03 | 0.04 | 0.06 | 0.04 |
| SD | 0.02 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 |

Table 16: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). NXY.TO
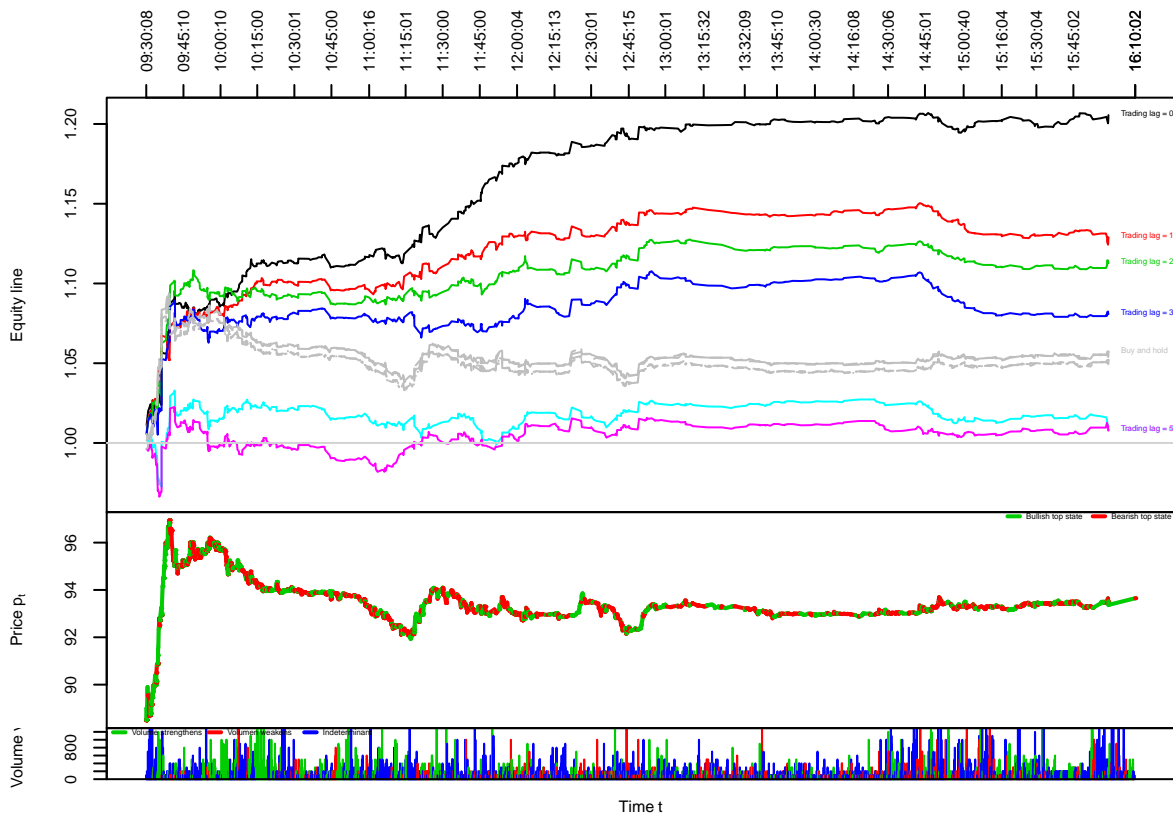


Figure 17: Out of sample equity line for NXY.TO [2007-05-09 09:30:00/2007-05-09 16:30:00]

### 5.2.9 SJRb.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | 0.00 | 0.01 | 0.01 | 0.03 | 0.02 | -0.00 | 0.01 |
| 2007-05-09 | -0.02 | 0.01 | 0.01 | -0.03 | -0.02 | -0.01 | 0.00 |
| 2007-05-10 | 0.00 | -0.04 | -0.07 | -0.06 | -0.04 | -0.02 | -0.06 |
| 2007-05-11 | 0.02 | -0.05 | -0.03 | -0.02 | -0.00 | -0.01 | 0.02 |
| 2007-05-14 | -0.02 | -0.02 | -0.02 | 0.01 | -0.00 | 0.00 | -0.01 |
| 2007-05-15 | -0.00 | 0.00 | -0.01 | 0.00 | -0.01 | -0.02 | -0.02 |
| 2007-05-16 | 0.02 | 0.01 | 0.01 | -0.01 | -0.01 | -0.01 | -0.02 |
| 2007-05-17 | 0.01 | 0.01 | -0.02 | -0.00 | 0.01 | -0.01 | 0.01 |
| 2007-05-18 | 0.00 | -0.05 | 0.02 | -0.01 | -0.01 | 0.01 | -0.01 |
| 2007-05-22 | -0.01 | -0.04 | -0.03 | -0.03 | -0.03 | -0.02 | -0.02 |
| 2007-05-23 | 0.01 | 0.02 | 0.01 | 0.04 | 0.03 | 0.01 | 0.01 |
| 2007-05-24 | -0.02 | -0.04 | -0.02 | -0.00 | -0.01 | -0.02 | -0.01 |
| 2007-05-25 | -0.00 | 0.01 | 0.03 | 0.02 | -0.01 | 0.00 | -0.00 |
| 2007-05-28 | -0.00 | 0.02 | 0.03 | 0.02 | 0.01 | 0.04 | 0.02 |
| 2007-05-29 | 0.01 | -0.02 | -0.01 | -0.05 | 0.00 | -0.00 | -0.01 |
| 2007-05-30 | 0.01 | 0.01 | 0.02 | 0.01 | 0.00 | 0.02 | 0.01 |
| 2007-05-31 | -0.01 | -0.03 | -0.02 | 0.01 | -0.01 | -0.01 | -0.03 |
| Total | 0.01 | -0.18 | -0.06 | -0.06 | -0.06 | -0.05 | -0.11 |
| Min | -0.02 | -0.05 | -0.07 | -0.06 | -0.04 | -0.02 | -0.06 |
| Mean | 0.00 | -0.01 | -0.00 | -0.00 | -0.00 | -0.00 | -0.01 |
| Median | 0.00 | 0.00 | -0.01 | -0.00 | -0.01 | -0.01 | -0.01 |
| Max | 0.02 | 0.02 | 0.03 | 0.04 | 0.03 | 0.04 | 0.02 |
| SD | 0.01 | 0.03 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 |

Table 17: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). SJRb.TO
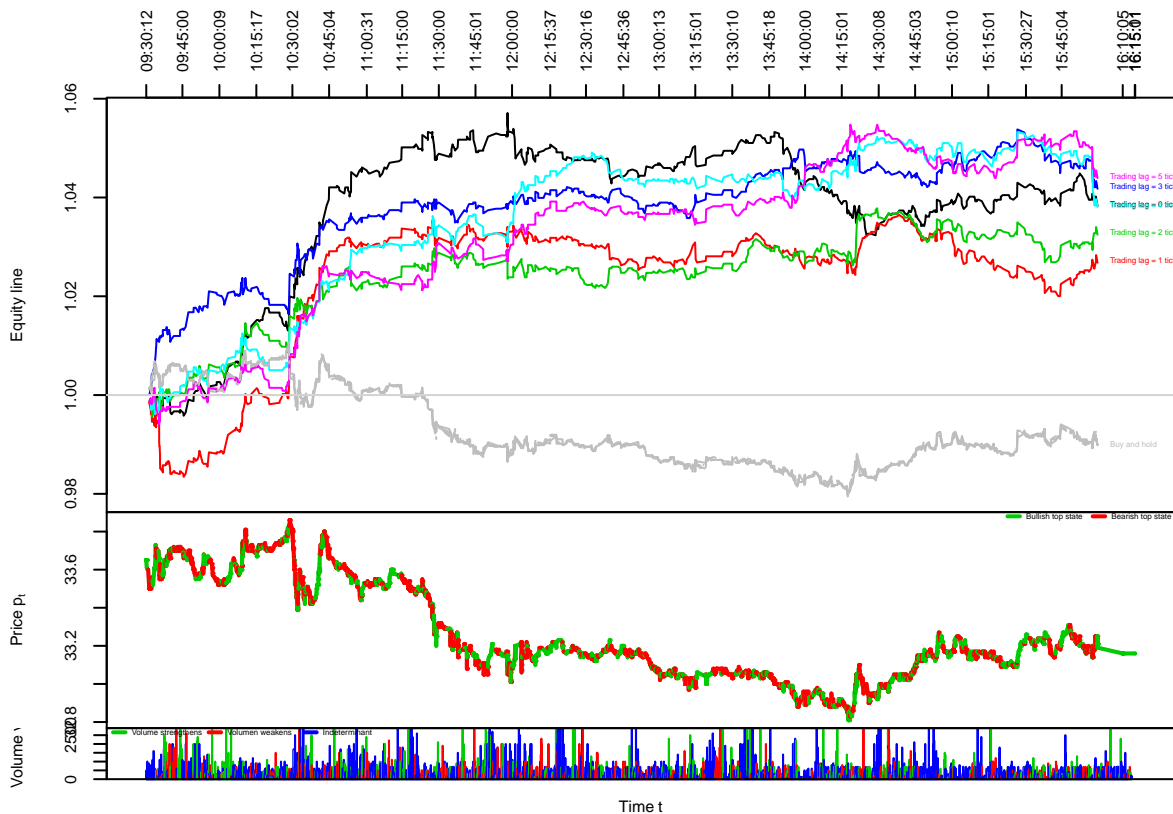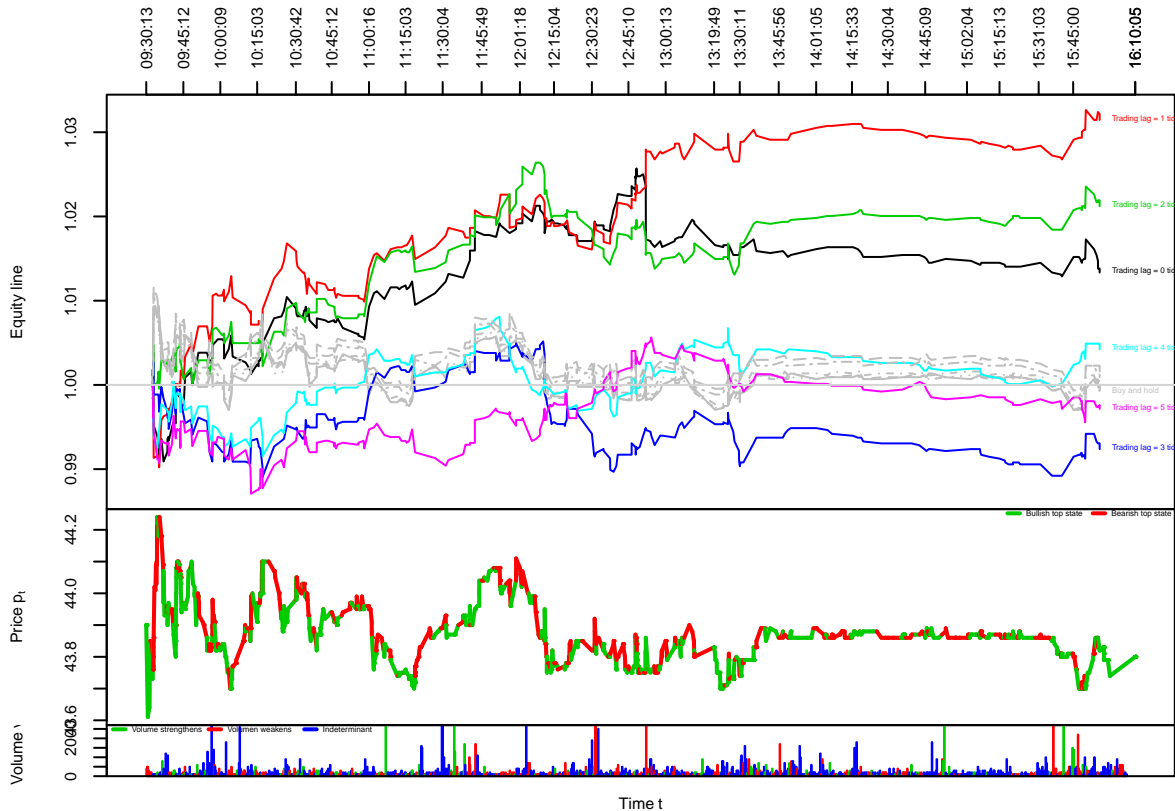


Figure 18: Out of sample equity line for SJRb.TO [2007-05-25 09:30:00/2007-05-25 16:30:00]

### 5.2.10 SU.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | 0.01 | 0.00 | 0.02 | 0.00 | 0.01 | 0.00 | 0.01 |
| 2007-05-09 | 0.01 | 0.04 | -0.02 | -0.01 | 0.01 | 0.01 | -0.01 |
| 2007-05-10 | -0.01 | 0.01 | 0.01 | 0.02 | 0.02 | -0.01 | -0.02 |
| 2007-05-11 | 0.02 | 0.02 | 0.02 | 0.02 | -0.00 | -0.01 | -0.01 |
| 2007-05-14 | -0.00 | 0.01 | -0.01 | -0.01 | -0.01 | -0.01 | 0.01 |
| 2007-05-15 | -0.02 | 0.02 | 0.01 | -0.02 | -0.01 | -0.02 | -0.02 |
| 2007-05-16 | 0.02 | 0.01 | 0.02 | 0.01 | -0.01 | 0.01 | 0.01 |
| 2007-05-17 | 0.03 | 0.04 | 0.04 | 0.03 | 0.02 | 0.02 | 0.03 |
| 2007-05-18 | 0.01 | 0.02 | -0.01 | -0.01 | -0.00 | -0.01 | -0.01 |
| 2007-05-22 | -0.00 | 0.05 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 |
| 2007-05-23 | 0.01 | -0.00 | 0.00 | 0.01 | 0.00 | 0.02 | 0.03 |
| 2007-05-24 | -0.03 | 0.04 | 0.04 | 0.03 | 0.02 | 0.04 | -0.01 |
| 2007-05-25 | 0.01 | 0.02 | 0.00 | 0.00 | -0.01 | 0.01 | 0.01 |
| 2007-05-28 | 0.00 | 0.01 | 0.01 | -0.01 | -0.01 | -0.00 | -0.01 |
| 2007-05-29 | -0.02 | 0.01 | -0.03 | -0.02 | -0.00 | -0.01 | -0.01 |
| 2007-05-30 | 0.01 | 0.04 | 0.02 | 0.01 | -0.02 | 0.00 | 0.01 |
| 2007-05-31 | 0.00 | 0.04 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 |
| Total | 0.03 | 0.45 | 0.18 | 0.08 | 0.04 | 0.09 | 0.05 |
| Min | -0.03 | -0.00 | -0.03 | -0.02 | -0.02 | -0.02 | -0.02 |
| Mean | 0.00 | 0.02 | 0.01 | 0.00 | 0.00 | 0.01 | 0.00 |
| Median | 0.01 | 0.02 | 0.01 | 0.01 | -0.00 | 0.00 | 0.01 |
| Max | 0.03 | 0.05 | 0.04 | 0.03 | 0.02 | 0.04 | 0.03 |
| SD | 0.01 | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.02 |

Table 18: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). SU.TO
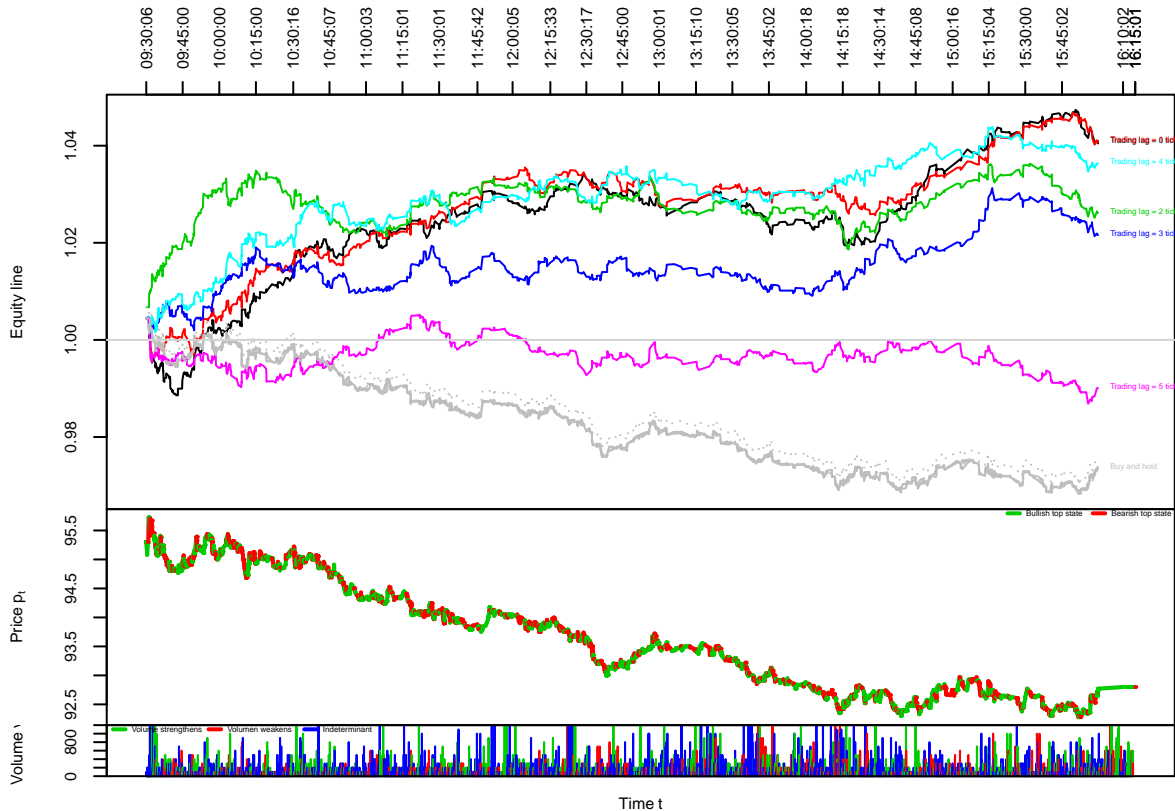


Figure 19: Out of sample equity line for SU.TO [2007-05-24 09:30:00/2007-05-24 16:30:00]

## 5.2.11 TCKb.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | -0.01 | -0.00 | -0.01 | -0.00 | -0.00 | 0.01 | 0.02 |
| 2007-05-09 | 0.02 | -0.00 | -0.02 | 0.01 | 0.04 | 0.04 | -0.00 |
| 2007-05-10 | 0.01 | 0.01 | 0.02 | 0.06 | 0.05 | 0.04 | 0.03 |
| 2007-05-11 | 0.03 | -0.02 | -0.00 | -0.01 | 0.01 | 0.01 | -0.02 |
| 2007-05-14 | -0.05 | 0.00 | -0.03 | -0.05 | -0.02 | -0.01 | -0.03 |
| 2007-05-15 | -0.00 | 0.05 | -0.01 | 0.01 | 0.03 | 0.03 | 0.03 |
| 2007-05-16 | -0.01 | -0.03 | -0.03 | -0.01 | -0.04 | -0.02 | 0.01 |
| 2007-05-17 | 0.00 | 0.08 | -0.03 | -0.04 | -0.03 | -0.03 | 0.00 |
| 2007-05-18 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.03 | -0.00 |
| 2007-05-22 | -0.04 | 0.01 | 0.00 | 0.02 | 0.02 | 0.02 | 0.02 |
| 2007-05-23 | -0.01 | -0.02 | -0.01 | -0.01 | -0.01 | 0.01 | -0.01 |
| 2007-05-24 | -0.04 | 0.06 | 0.04 | 0.04 | 0.02 | 0.02 | 0.04 |
| 2007-05-25 | 0.01 | 0.02 | -0.02 | -0.02 | -0.03 | -0.01 | -0.01 |
| 2007-05-28 | 0.01 | -0.01 | -0.01 | 0.01 | 0.01 | -0.00 | 0.00 |
| 2007-05-29 | -0.02 | -0.00 | -0.02 | -0.01 | -0.00 | -0.00 | 0.00 |
| 2007-05-30 | 0.04 | 0.09 | 0.05 | 0.07 | 0.06 | 0.08 | 0.05 |
| 2007-05-31 | 0.00 | 0.02 | 0.01 | 0.03 | -0.01 | 0.00 | -0.01 |
| | | | | | | | |
| Total | -0.06 | 0.29 | -0.06 | 0.11 | 0.10 | 0.23 | 0.12 |
| Min | -0.05 | -0.03 | -0.03 | -0.05 | -0.04 | -0.03 | -0.03 |
| Mean | -0.00 | 0.02 | -0.00 | 0.01 | 0.01 | 0.01 | 0.01 |
| Median | 0.00 | 0.01 | -0.01 | 0.01 | 0.01 | 0.01 | 0.00 |
| Max | 0.04 | 0.09 | 0.05 | 0.07 | 0.06 | 0.08 | 0.05 |
| SD | 0.02 | 0.03 | 0.02 | 0.03 | 0.03 | 0.03 | 0.02 |

Table 19: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). TCKb.TO
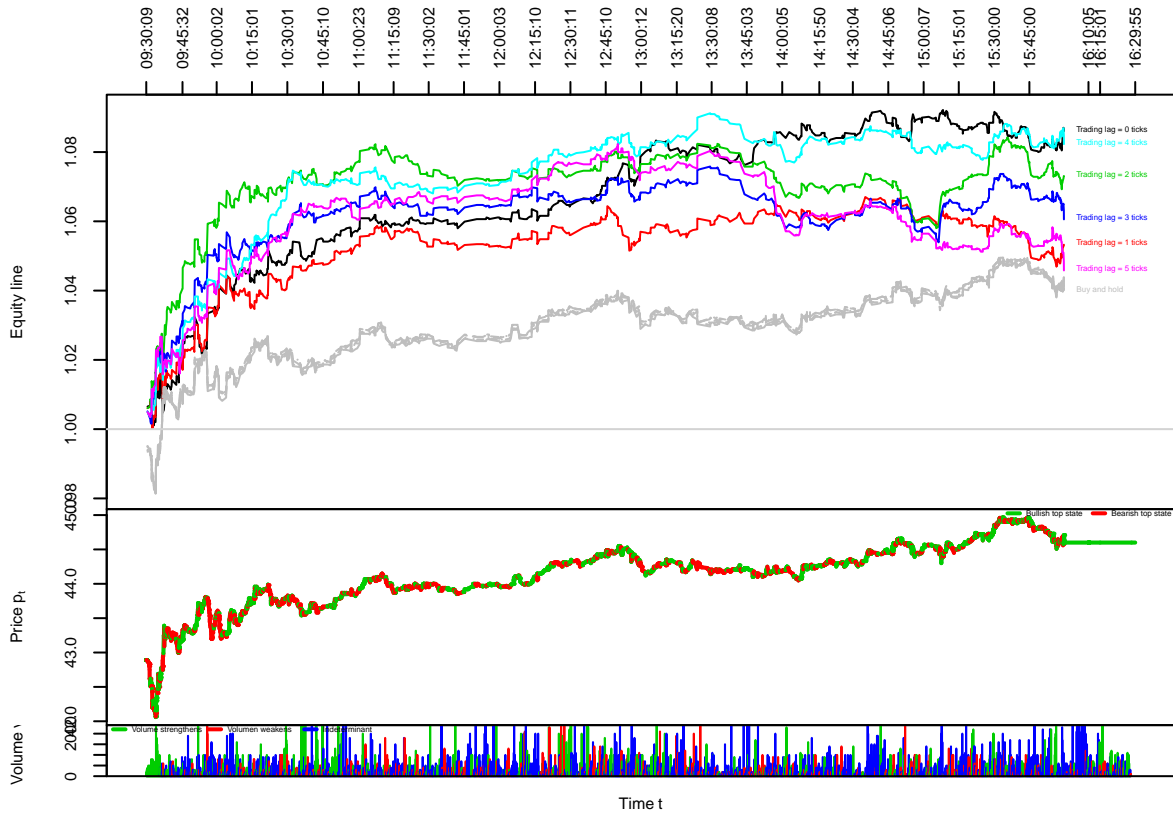


Figure 20: Out of sample equity line for TCKb.TO [2007-05-30 09:30:00/2007-05-30 16:30:00]

### 5.2.12 TLM.TO

| | Buy-Hold | HHMM (0 lag) | HHMM (1 lag) | HHMM (2 lag) | HHMM (3 lag) | HHMM (4 lag) | HHMM (5 lag) |
|---|---|---|---|---|---|---|---|
| 2007-05-08 | 0.01 | 0.03 | -0.00 | -0.02 | -0.01 | -0.01 | -0.01 |
| 2007-05-09 | -0.03 | 0.01 | -0.00 | -0.01 | 0.00 | 0.03 | 0.01 |
| 2007-05-10 | -0.02 | 0.00 | -0.01 | -0.02 | 0.00 | -0.00 | 0.01 |
| 2007-05-11 | 0.03 | -0.02 | -0.03 | -0.01 | -0.01 | -0.00 | -0.00 |
| 2007-05-14 | -0.00 | -0.00 | 0.01 | -0.01 | 0.01 | 0.00 | -0.02 |
| 2007-05-15 | -0.01 | -0.01 | -0.01 | -0.03 | -0.01 | -0.01 | -0.01 |
| 2007-05-16 | 0.01 | 0.01 | 0.02 | 0.01 | 0.02 | 0.02 | 0.00 |
| 2007-05-17 | 0.02 | -0.01 | 0.01 | -0.01 | -0.01 | -0.01 | -0.00 |
| 2007-05-18 | -0.00 | -0.05 | -0.01 | -0.01 | -0.01 | -0.01 | 0.01 |
| 2007-05-22 | -0.01 | -0.01 | -0.02 | -0.02 | -0.02 | -0.02 | -0.01 |
| 2007-05-23 | 0.01 | 0.01 | -0.02 | -0.02 | -0.00 | 0.02 | 0.01 |
| 2007-05-24 | -0.02 | -0.00 | 0.02 | 0.02 | 0.01 | 0.02 | 0.04 |
| 2007-05-25 | 0.00 | -0.02 | -0.00 | 0.02 | 0.01 | 0.01 | 0.01 |
| 2007-05-28 | 0.01 | -0.02 | -0.01 | 0.00 | -0.01 | 0.00 | -0.01 |
| 2007-05-29 | -0.02 | 0.01 | 0.01 | 0.02 | -0.02 | 0.00 | 0.01 |
| 2007-05-30 | 0.04 | -0.01 | 0.01 | 0.03 | 0.01 | -0.01 | 0.01 |
| 2007-05-31 | -0.02 | -0.02 | 0.01 | -0.03 | -0.02 | 0.00 | 0.00 |
| | | | | | | | |
| Total | -0.02 | -0.09 | -0.03 | -0.10 | -0.05 | 0.04 | 0.07 |
| Min | -0.03 | -0.05 | -0.03 | -0.03 | -0.02 | -0.02 | -0.02 |
| Mean | -0.00 | -0.01 | -0.00 | -0.01 | -0.00 | 0.00 | 0.00 |
| Median | -0.00 | -0.01 | -0.00 | -0.01 | -0.01 | 0.00 | 0.00 |
| Max | 0.04 | 0.03 | 0.02 | 0.03 | 0.02 | 0.03 | 0.04 |
| SD | 0.02 | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 |

Table 20: Compound daily return originated in the HHMM trading strategy for different levels of lags. Lag measured in ticks between the end of the zig-zag and the execution of the trade (zero lag suffers from look-ahead bias). TLM.TO
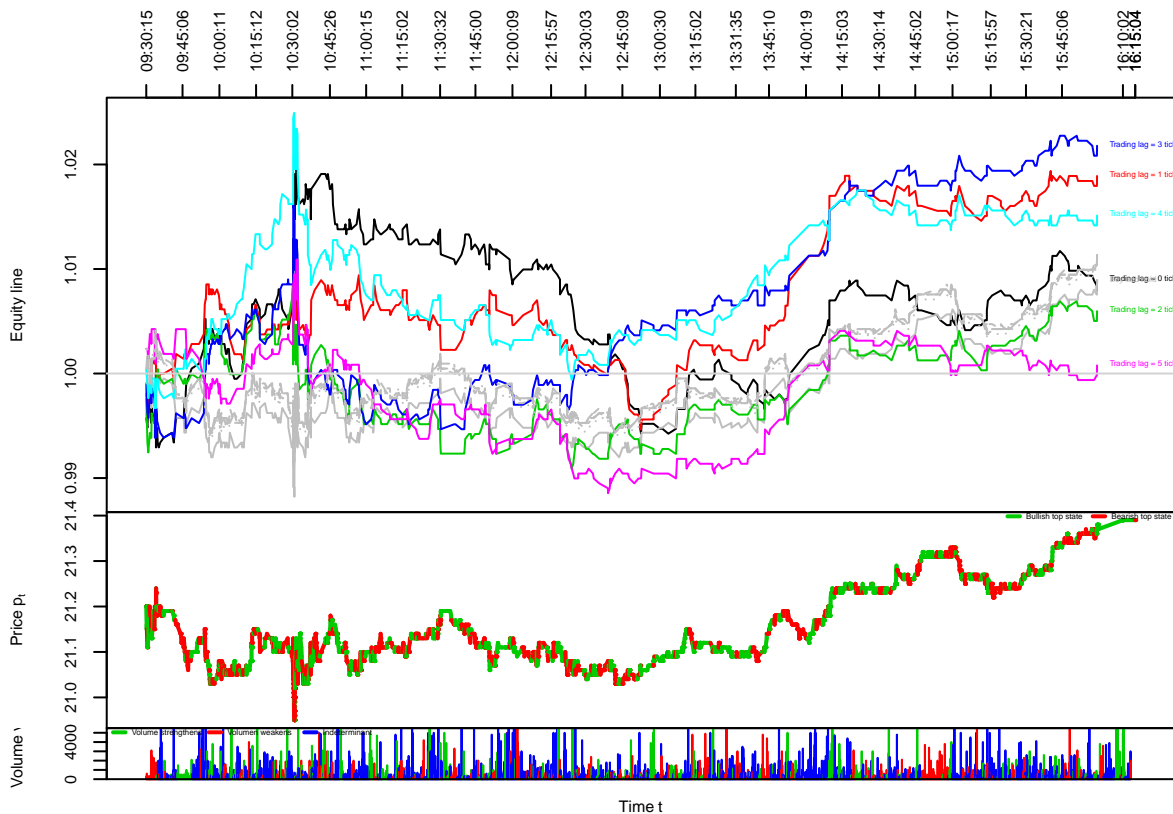


Figure 21: Out of sample equity line for TLM.TO [2007-05-16 09:30:00/2007-05-16 16:30:00]

## 5.3   Original Computing Environment

```
## 
## CXXFLAGS=-O3 -mtune=native -march=native -Wno-unused-variable -Wno-unused-function

## Session info ----------------------------------------------------------------

##   setting  value
##   version  R version 3.3.3 (2017-03-06)
##   system   x86_64, mingw32
##   ui       RTerm
##   language (EN)
##   collate  Spanish_Argentina.1252
##   tz       America/Buenos_Aires
##   date     2017-08-29

## Packages --------------------------------------------------------------------

##   package      * version    date       source
##   BH             1.62.0-1   2016-11-19 CRAN (R 3.3.2)
##   colorspace     1.3-2      2016-12-14 CRAN (R 3.3.3)
##   dichromat      2.0-0      2013-01-24 CRAN (R 3.3.2)
##   digest       * 0.6.12     2017-01-27 CRAN (R 3.3.3)
##   ggplot2      * 2.2.1      2016-12-30 CRAN (R 3.3.3)
##   graphics     * 3.3.3      2017-03-06 local
##   grDevices    * 3.3.3      2017-03-06 local
##   grid           3.3.3      2017-03-06 local
##   gridExtra      2.2.1      2016-02-29 CRAN (R 3.3.3)
##   gtable         0.2.0      2016-02-26 CRAN (R 3.3.3)
##   inline         0.3.14     2015-04-13 CRAN (R 3.3.3)
##   labeling       0.3        2014-08-23 CRAN (R 3.3.2)
##   lattice        0.20-34    2016-09-06 CRAN (R 3.3.3)
##   lazyeval       0.2.0      2016-06-12 CRAN (R 3.3.3)
##   magrittr       1.5        2014-11-22 CRAN (R 3.3.3)
##   MASS           7.3-45     2016-04-21 CRAN (R 3.3.3)
##   Matrix         1.2-8      2017-01-20 CRAN (R 3.3.3)
##   methods      * 3.3.3      2017-03-06 local
##   munsell        0.4.3      2016-02-13 CRAN (R 3.3.3)
##   plyr           1.8.4      2016-06-08 CRAN (R 3.3.3)
##   RColorBrewer   1.1-2      2014-12-07 CRAN (R 3.3.2)
##   Rcpp           0.12.10    2017-03-19 CRAN (R 3.3.3)
##   RcppEigen      0.3.2.9.1  2017-03-15 CRAN (R 3.3.3)
##   reshape2       1.4.2      2016-10-22 CRAN (R 3.3.3)
##   rstan        * 2.14.2     2017-03-19 CRAN (R 3.3.3)
##   scales         0.4.1      2016-11-09 CRAN (R 3.3.3)
##   StanHeaders  * 2.14.0-1   2017-01-09 CRAN (R 3.3.3)
##   stats        * 3.3.3      2017-03-06 local
##   stats4         3.3.3      2017-03-06 local
##   stringi        1.1.3      2017-03-21 CRAN (R 3.3.3)
##   stringr        1.2.0      2017-02-18 CRAN (R 3.3.3)
##   tibble         1.3.0      2017-04-01 CRAN (R 3.3.3)
##   tools          3.3.3      2017-03-06 local
##   utils        * 3.3.3      2017-03-06 local
```