Machine Learning, Technical Analysis and Algorithmic Trading

Source: Germán Creamer and Yoav Freund (2007). “A Boosting Approach for Automated Trading.” Journal of Trading 2 (3): 84-96.

Trading requires finding patterns in massive..

amounts of information in (semi) automated way → Machine Learning

Objective of this presentation:

- Review price discovery process
- Overview of cutting-edge developments in machine learning and how financial markets are applying these principles
- Deeper exploration of a machine learning method that may lead to uncovering profit opportunities
- Practical tips for applying machine learning techniques to futures trading
- How technique work in practice

What is Behind

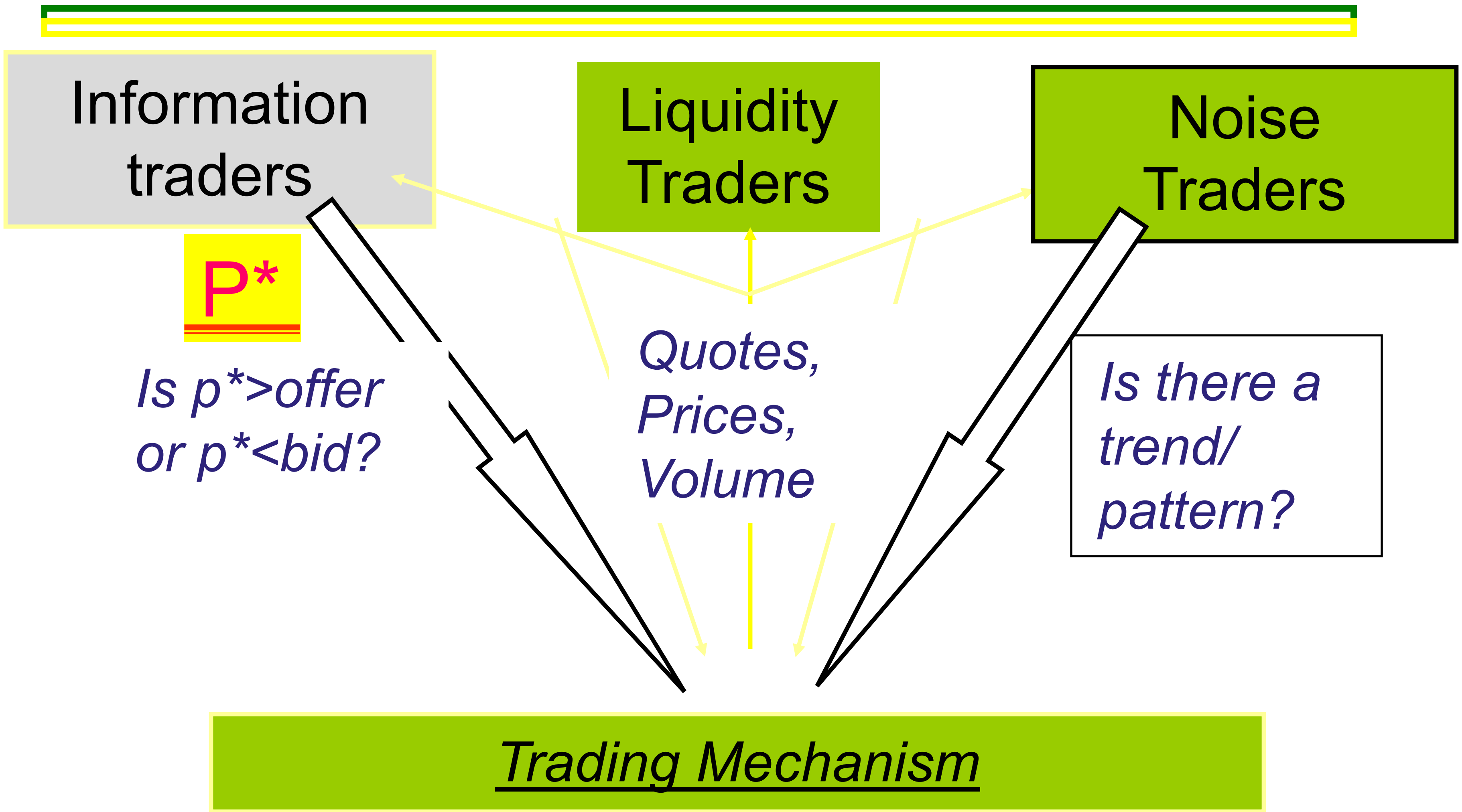
Standard Finance Models?

- Markets are informationally efficient (EMH)
- Shares have unique fundamental values
- Informed investors form identical expectations
    - Homogeneous expectations

The Standard Microstructure Model

- Information traders
- Liquidity traders
- Noise traders

Orders From 3 Types of Participants



What motivates individuals to trade?

Accepted academic answer

- Informed traders
- Liquidity traders
- Noise traders

Perhaps we should add a fourth

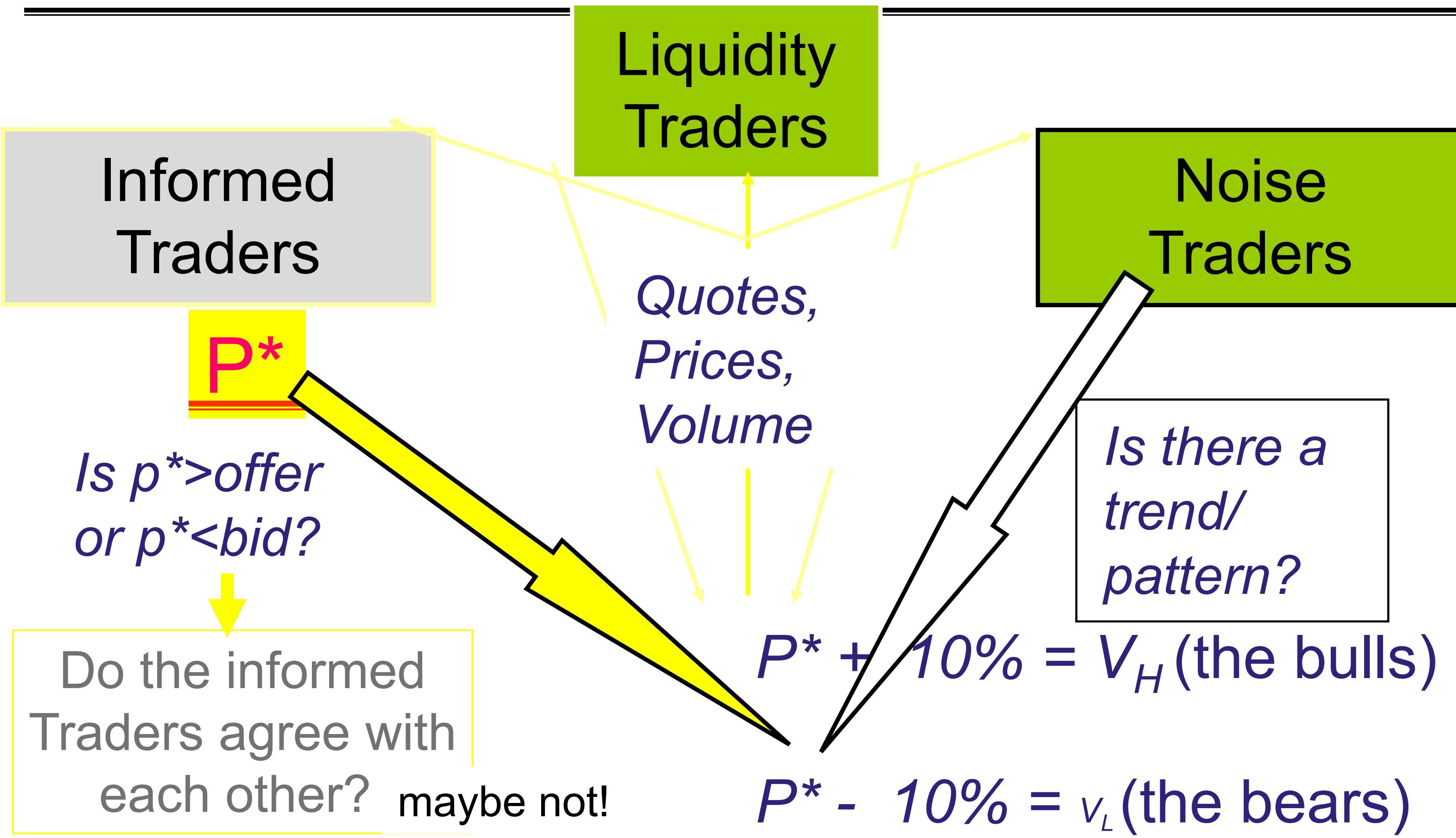
- Divergent expectations

(people disagree…)

Divergent Expectations Has Implications For

- Understanding market structure & operations
- Assessing market quality
- Government regulatory policy
- Understanding volatility

Representing Divergent Expectations



The Inside Scoop on Price Discovery

- A complex, protracted process
- Contributes to intra-day volatility
- Equilibrium depends on the sequence of order arrivals & on how orders are handled

    → A coordination problem

- The quality of price discovery depends on trader behavior & market structure
- Divergent expectations underlie the complexity of price discovery

Divergent Expectations:A Simple Setting

- A company is facing a jury trial – its share value will be affected appreciably by the outcome
- Investors can have 1 of 2 expectations
  - Some believe pr(acquittal) = .80
  - Some believe pr(acquittal) = .35
- Shares are valued at
  - $55 by those who expect acquittal
  - $45 by those who expect conviction

Use data driven methods (such as machine learning) to infer the divergent expectations of economic agents.

Machine learning (ML) applications in finance

- Microstructure and algorithmic trading:
  - Replication of technical analysis and time series patterns
  - Generation of new trading rules & optimal combination of existent trading rules
- Portfolio optimization:
  - Use ML to propose financial scenarios (investors' view) in Black Litterman model
- Equity valuation and fundamental analysis:
  - Decision trees or Bayes Nets to define main drivers of corporate profitability and equity value
- Financial forecasting: using financial time series with unstructured data (images, text [news, tweets…] to forecast financial trends
  - Anticipate major events at the macro or at the customer level
- Risk management: Credit rating, default probability, volatility and customer segmentation
- Consumer analytics:
  - Managing customer data: extract behavioral information from customers' decisions and transactions
    - Social networks & access to new investors through current customers
- Personalization and customization: get to know the customer to offer the best product/advice.
- Fraud detection

Technical analysis detects and interprets patterns…

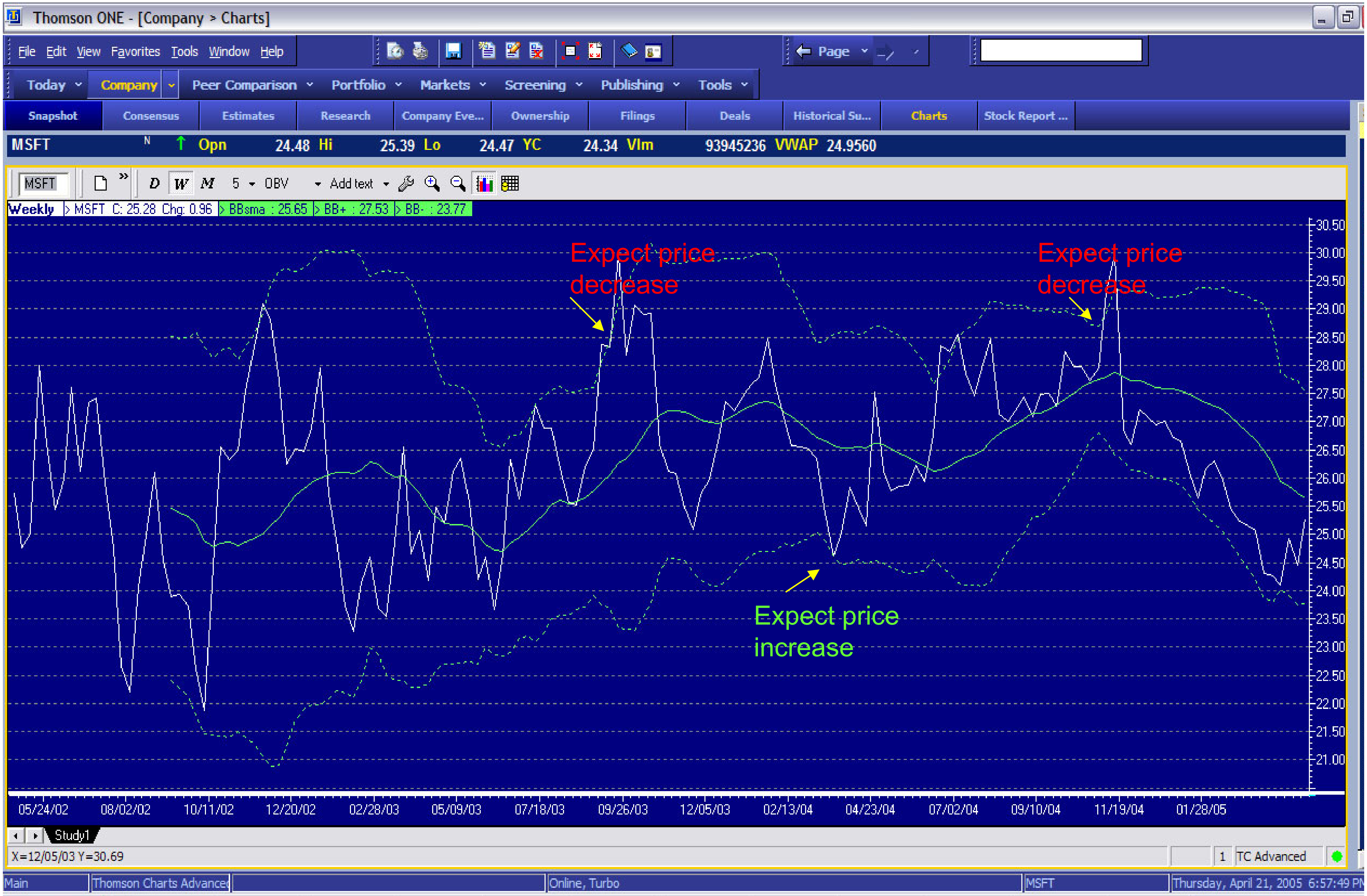
…in past futures prices based on charts to make future investment decisions

- It shares with machine learning the interest on pattern recognition
- Formalizes traders' thinking:
  - "buy when it breaks through a high"
  - "sell when it is declining", etc.
- Type of indicators:
  - Price
  - Momentum and oscillation indicators
  - Volatility indicators
  - Volume indicators…

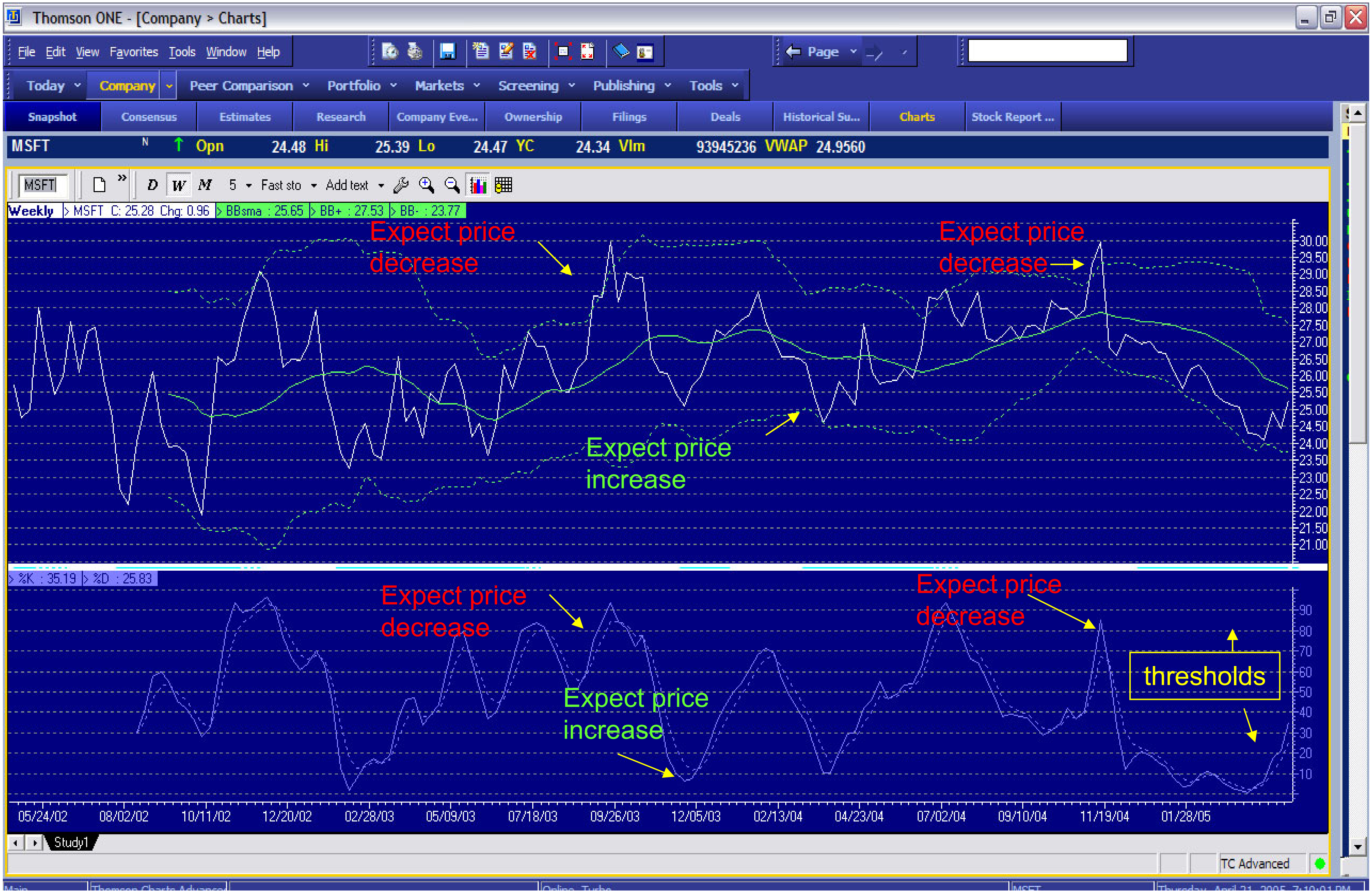
- Price moving average:
  - Simple
  - Exponential
- Bollinger bands

Moving Averages: Mean of last n observations of a time series

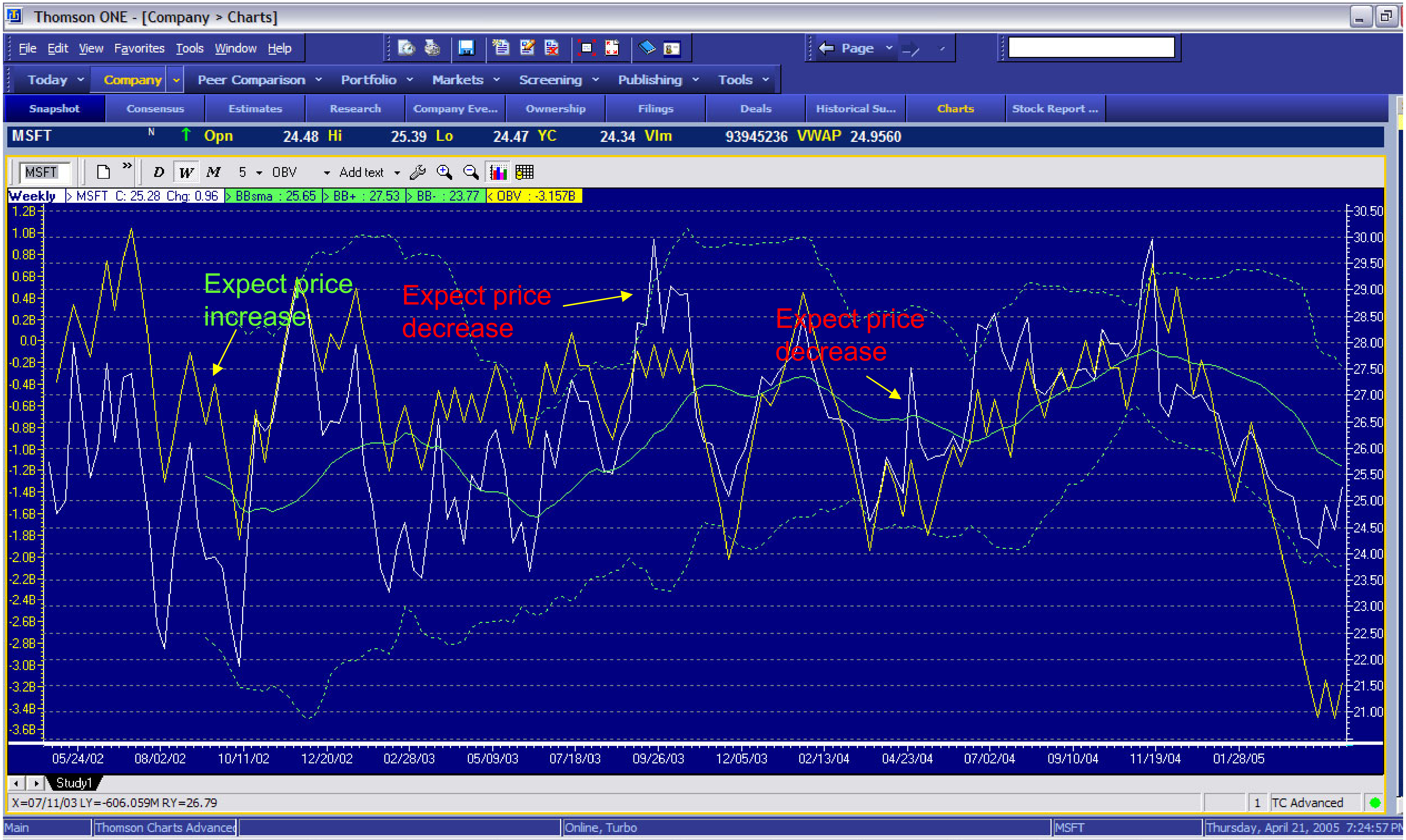
Bollinger Bands: moving average +/- 2 st.dev.



Bollinger bands + Stochastic indicator



On balance volume: adds (reduce) volume with price increase (decrease)

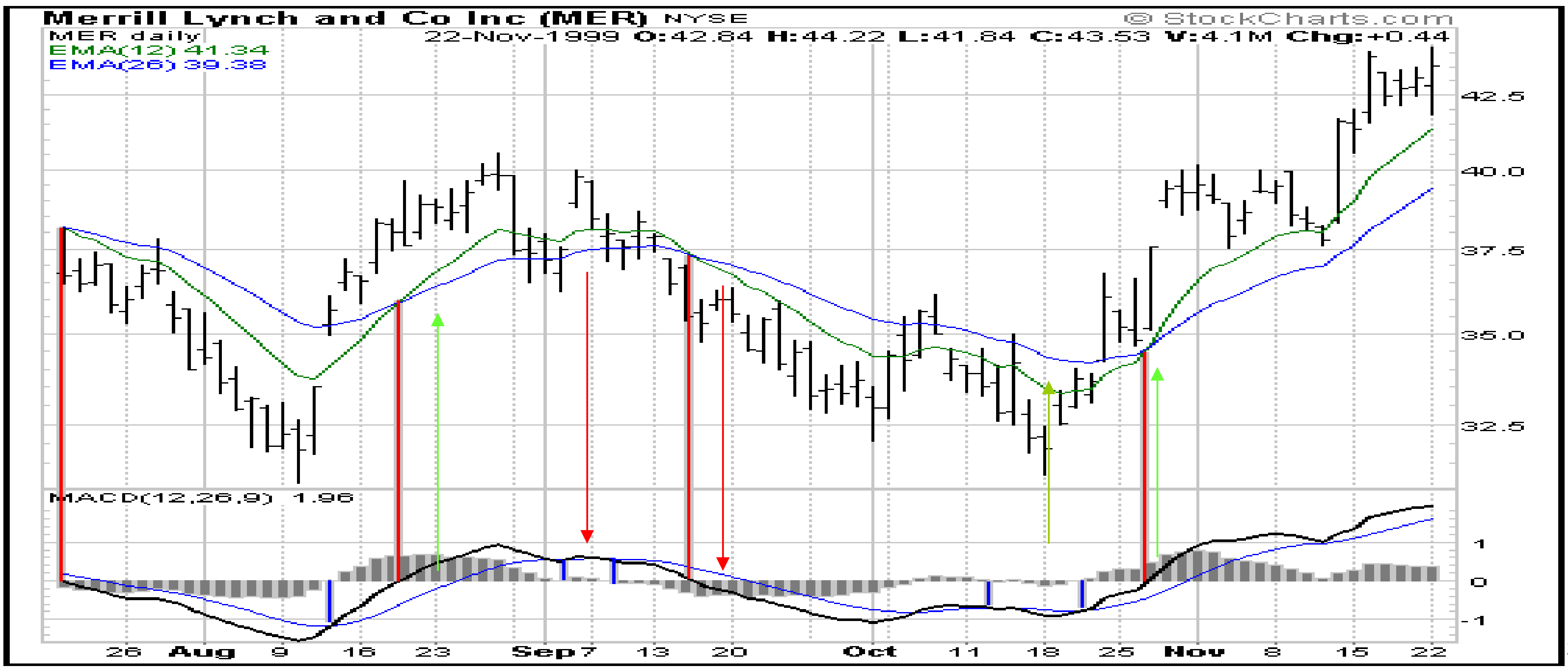


Momentum & oscillation indicators

- Momentum
- Acceleration
- Rate of change
- Moving average convergence divergence
- Relative strength index
- Stochastic oscillator
- Williams indicator
- Money flow index

Moving average converge divergence (MACD)

- Difference between 2 moving average of periods
- Price change signal when crosses above (below) signal line



Volatility indicators:

- Chaikin volatility
- Garman Klass volatility
- GARCH

Volatility Indicators: Autoregressive Conditionally Heteroscedastic (ARCH)

Engle 1982

- simulate serial correlation of volatility of a stochastic process (log-return: r_1, r_2, …)

$$\sigma_t^2(n) = \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + \alpha_m a_{t-m}^2$$

where:

- $a_t = r_t - \mu$ (mean adjusted return)
- $r_t = \log(p_{t+1} / p_t)$
- □ μ as the mean return
- □ $\alpha_0 > 0$ and α_i ¸ 0
- □ $\varepsilon_1, \varepsilon_2, \ldots,$ is a random sequence of "noise" random variables according to N(0,1).

Generalized Autoregressive Conditionally Heteroskedastic (GARCH)

Bollerslev 1986

- simulate volatility without having to calculate a large number of coefficients for polynomials of high-order.
- r_t can be simulated with an autoregressive moving-average (ARMA) model.
- GARCH(m,s) adds a distributed lag structure to simulate the conditional variance:

$$\sigma_t^2(n) = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{i=1}^{s} \beta_j \sigma_{t-j}^2(n)$$

Volume indicators:

–On balance volume

–Accumulation/distribution line

–Chaikin oscillator

–Negative/Positive volume index

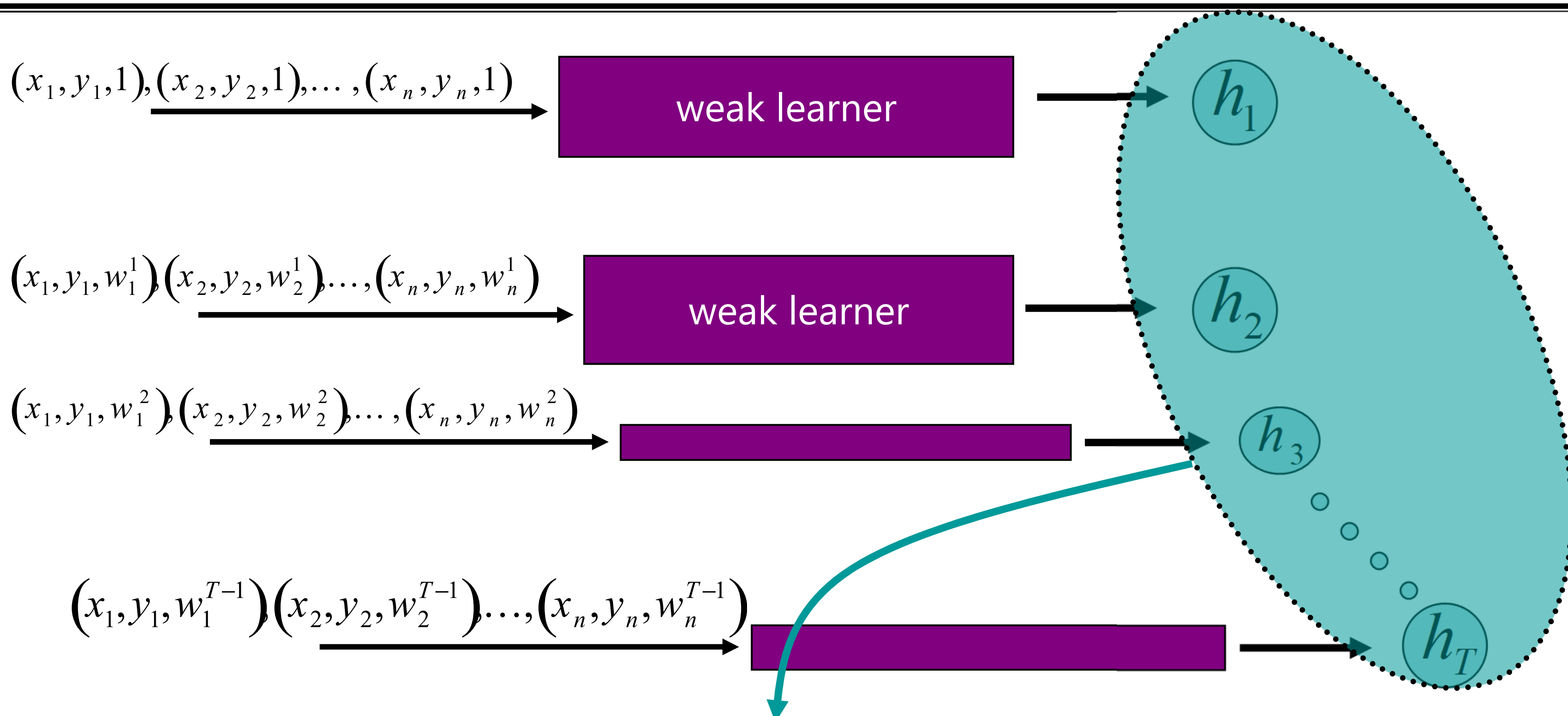
–Price – volume trend

Boosting: a method for improving classifier accuracy

Freund, Schapire 1997

- Basic idea:
  - Perform iterative search to locate the regions/ examples that are more difficult to predict.
  - Through each iteration reward accurate predictions on those regions.
  - Combines the rules from each iteration.
- Only requires that the underlying learning algorithm be better than guessing.

The boosting process: General discriminative learning algorithm (Freund & Schapire, 1997)

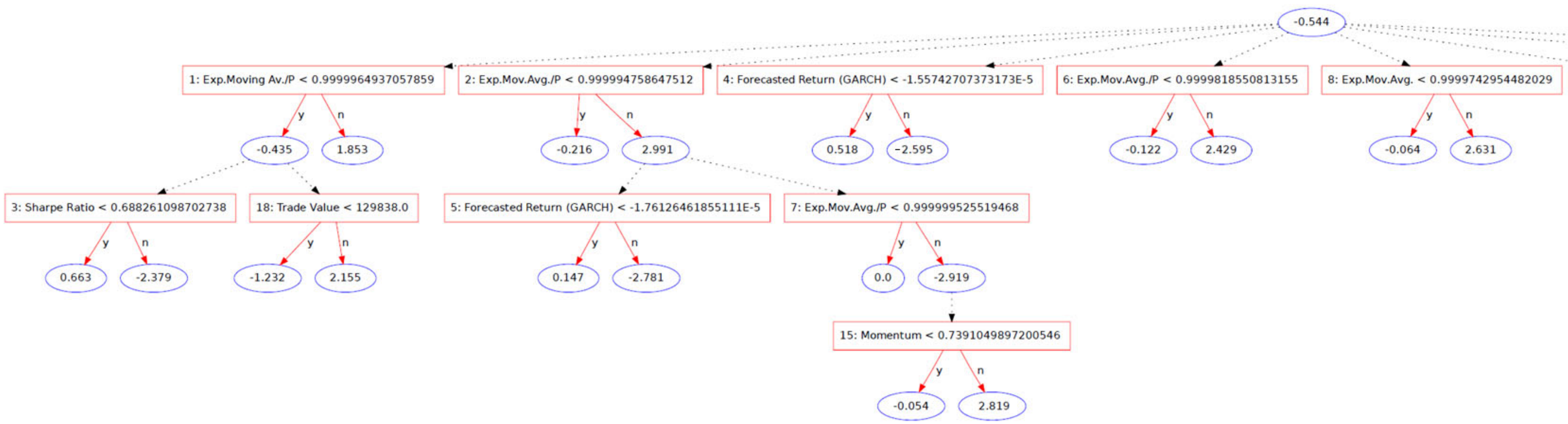


$$F_T(x) = \alpha_1 h_1(x) + \alpha_2 h_2(x) + \cdots + \alpha_T h_T(x)$$

Final rule: $f_T(x) = \mathrm{sign}(F_T(x))$

Logitboost is a modification of Adaboost and is an algorithm for step-wise logistic regression (Friedman et al. (2000)).

Rules of boosting: alternating decision tree (ADT) for investment decisions



ADT for EURO STOXX 50 ® Index Futures (FESX), 03-04-2009

Trading alg. – Boosting (Creamer & Freund, 2007):

Artificial trader CRP_TA combines:

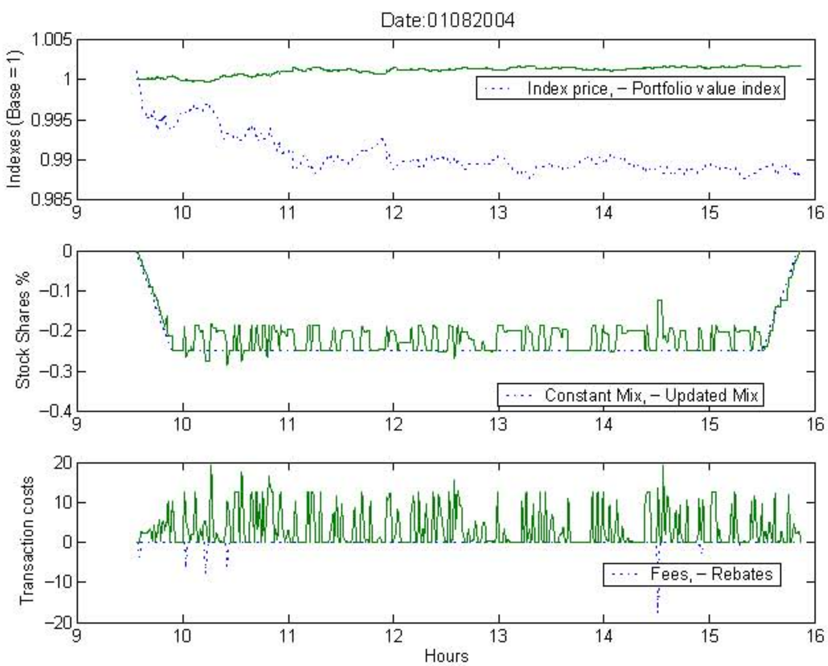
- 1) Logitboost: forecasts the direction of the stock price using ADTs which are implemented with Logitboost.
  - Training using technical analysis indicators (90 days).
  - Logitboost combines technical indicators and generates a new set of trading rules based on the market conditions.
  - According to daily market forecast, algorithm establishes a long position (50% of portfolio invested in asset), short (25% of portfolio) or hold (q_g).
- 2) Constant rebalanced portfolio:
  - Starts with a balanced position according to the % shares over portfolio value established as a goal based on logitboost forecast.
  - Sends simultaneously a buy and sell limit order:
    - If $q_t < q_g - \Delta / W$, then send buy limit order for Δ.
    - If $q_t > q_g + \Delta / W$, then send sell limit order for Δ.
    - Otherwise, hold.

  where:
  - q_g: goal mix of stocks and cash
  - q_t: current mix of stocks and cash
  - W : net value portfolio
  - Δ : order size (in dollars)

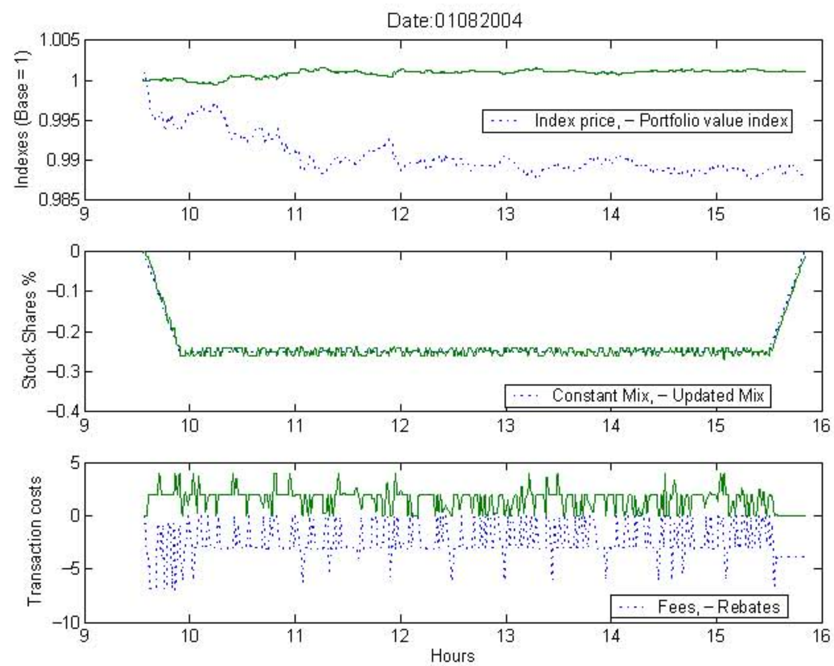
Source: Germán Creamer and Yoav Freund (2007). "A Boosting Approach for Automated Trading." Journal of Trading 2 (3): 84-96

Improvement of CRP_TA with a market maker strategy (market maker CRP_TA) and tested during the period Jan.5-9,2004:

- Starts with a balanced position according to the % shares over portfolio value established as a goal (q_g) based on the logitboost forecast.
- Sends simultaneously:
  - a buy limit order at price slightly below (0.005) than price at top of the buy order book
  - a sell limit order at price slightly above (0.005) than price at top of the sell order book
- If the order is not 100% filled within ten minutes of being issued, existent limit orders are cancelled, and limit orders are reissued.
- Results are partially explained as the Exchange pays a rebate for limit orders and charges a higher fee for market order. This strategy eliminates or minimizes market order, so it benefits from the adequate forecast of market direction and for the accumulation of rebates.



(a) Market maker CRP_TA



(b) CRP_TA

ML tips for trading

- Increasing research on ML and trading: both equity trading and index futures trading have a similar forecasting problem: anticipate price trend
- For trading, more important to forecast trend (classification problem: up, down, same) than an exact price (regression problem)
- Combination of trading rules improve results vs. application of a single trading rule
- Genetic algorithm or decision trees (ADTs) generate trading rules that are interpretable while “Black box” algorithms (i.e. neural networks) generate uninterpretable rules: clarity of trading rules improve acceptance & innovation by traders
- Select adequate performance indicator: machine learning methods are typically evaluated using test error (forecast vs. actual), however evaluation of trading strategies should emphasize risk adjusted return measures (i.e. Sharpe ratio)

ML tips for trading

- Most of the ML research applied to trading is based on daily prices and the objective is to forecast trend of next day. Moving to high-frequency data may increase variance, so forecasts must be interpreted accordingly.
  - Dempster, Bates, and Romahi (2003) show that for FX trading using order flow and order book data is usually superior that trading on technical signal alone
- Tests become more demanding as forecasting methods improve: need of continuous innovation and testing
- Ensemble (bagging, boosting, random forests) and hybrid methods (i.e. neuro-genetic algorithm) may have better performance than individual methods as they can be used for different objectives. For example, use boosting for prediction and genetic algorithm to calibrate the model

ML tips for trading

- Pre-processing data based on:
  - Feature selection
  - Features transformation:
    - Using expert or domain knowledge
    - Meaningful ratios. Example: Price / Earnings brings more information than Price and Earnings as two independent features
  - Normalization/standardization (if required)
  - Data smoothing (simple or exponential moving average)
- Establish minimum level of quality/strength of predictor to take action and strict risk management rules to avoid major jumps:
  - Some learning algorithms that are very good predictors may show high variance with small changes in the sample (bias vs. variance reduction conflict)