

## Unexpected Features of Financial Time Series: Higher-Order Anomalies and Predictability

Erhard Reschenhofer  
*University of Vienna*

*Abstract:* Examining the daily Dow Jones Industrial Average (DJI) we find evidence both of higher-order anomalies and predictability. While most researchers are only aware of the relatively harmless anomalies that occur just in the mean, the first part of this article provides empirical evidence of more dangerous kinds of anomalies occurring in higher-order moments. This evidence casts some doubt on the common practice of fitting standard time series models (e.g., ARMA models, GARCH models, or stochastic volatility models) to financial time series and carrying out tests based upon autocorrelation coefficients without making proper provision for these anomalies. The second part of this article provides evidence in favor of the predictability of the returns on the DJI and, more interestingly, against the efficient market hypothesis. The special value of this evidence is due to the simplicity of the involved methods.

*Key words:* Day-of-the-week effects, efficient market hypothesis, nonstationarities.

### 1. Introduction

It is well known that stock prices do not follow a pure random walk. Price changes are neither independent nor identically distributed. Firstly, there are various linear and nonlinear dependencies between successive price changes, e.g., serial correlation, periodic patterns, and conditional heteroskedasticity. Secondly, both the unconditional distribution of the price changes and the conditional distributions (dependencies) change over time. Any test for a particular deviation from randomness must therefore be designed in such a way that it takes all the other possible deviations into account. Unfortunately, that is easier said than done. For example, the tests that are typically used to detect serial correlation (see, e.g., Taylor, 1984, Lo and MacKinlay, 1988, and Deo, 2000) are only robust against conditional heteroskedasticity but not against other peculiarities of financial time series like nonstationarities and calendar effects. The researchers using these tests tend to play down this disadvantage. They argue that it is always possible to divide a nonstationary series of returns into roughly stationary

segments and that tests based upon autocorrelation coefficients need not be affected by calendar effects like the January effect and the weekend effect because these effects occur only in the mean and discrepancies in the mean are small compared to the variance. However, the empirical evidence presented in Section 2 casts some doubt on these arguments. Multiple nonstationarities are described in Subsection 2.1, followed in Subsection 2.2 by a description of calendar effects in higher-order moments. Concluding remarks occur at the end of each section.

In Section 3, we examine the efficient market hypothesis. A crucial question in financial economics is whether or not there is a strategy for buying and selling stocks that yields superior returns than the buy-and-hold strategy. Clearly, it is practically impossible to rule out that such a strategy exists because there are too many possibilities. So if this question cannot be answered positively, it must remain unanswered. Amazingly, the majority of researchers dealing with this problem do not first and foremost try to construct profitable trading strategies but rather focus on a related question, namely whether or not stock returns are predictable (early exceptions to this rule are Alexander, 1961, Fama and Blume, 1966, Levy, 1967, and Stevenson and Bear, 1970, Leuthold, 1972; more recent ones are Taylor, 1986, Jegadeesh and Titman, 1993, Pesaran and Timmerman, 1995, Gencay, 1998, Qi, 1999, and Sullivan et al., 1999). A natural approach to test whether price changes are predictable is to use past prices and additional economic variables to construct a nontrivial predictor for future price changes and to test whether the sum of squared prediction errors implied by this predictor is smaller than that implied by the trivial predictor, which is given by the mean of the past price changes. Of course, we must impose suitable restrictions on the price changes in order to make this testing problem tractable. In the simplest case, where the nontrivial predictor is a function of past price changes only, some robust version of the random walk hypothesis (allowing, e.g., for a nonzero mean and conditional heteroskedasticity) could be used as null hypothesis. A severe shortcoming of this approach is that statistically significant predictability is not necessarily of speculative value, particularly when transaction costs are taken into account. Another disadvantage is that the tests that are typically used to detect predictability are not robust against the anomalies described in Section 2.

Fortunately, we do not really need a significance test to furnish proof of the predictability of the returns. Using a very simple trading strategy, Alexander (1961) succeeded already long time ago in proving predictability. According to Alexander's filter rule the risky asset is purchased when its price increases by  $v\%$  and (short) sold when its price drops by  $v\%$ . In the absence of transaction costs, the filter rule beats the buy-and-hold strategy for small values of  $v$ . But transaction costs typically eliminate the profits (see Fama and Blume, 1966). This is particularly true for very small values of  $v$ . On the other hand, positive

---

results obtained with values of  $v$  as large as 5% or even 10% (Stevenson and Bear, 1970, Leuthold, 1972) must be treated with caution because in these cases the number of trades is relatively small and hence the overall positive result could have been obtained just by luckily avoiding a few sharp falls in prices. Moreover, there is the problem of retrospective optimization. How many different trading rules had been examined before the “profitable” rule was found?

The results obtained with alternative trading strategies are hardly more reliable. For example, Levy (1967) came up with his trading strategy of buying stocks with current prices that are substantially higher than their average prices over the past 27 weeks after he had examined 68 different trading rules in his dissertation (see Jensen and Bennington, 1970). Moreover, Levy’s strategy as well as other strategies that buy stocks that have performed well in the past and sell stocks that have performed poorly in the past (relative strength strategies) conflict with the popular view that stock prices overreact to new information. If this view were correct, one would rather expect that the opposite strategy is profitable. Indeed, for very short time horizons of one week or one month, Jegadeesh (1990) and Lehmann (1990) reported abnormal returns generated by selling past winners and buying past losers. However, after realizing that practitioners used the original relative strength strategy for time horizons of 3 to 12 months, Jegadeesh and Titman (1993) examined this strategy just for these time horizons and found “significant” positive returns. In this case, the possibility of a selection bias is due to the fact that Jegadeesh and Titman used the same data as the practitioners. An additional problem of relative strength strategies is that they switch between different risky assets. Any extra profit could therefore be merely due to a higher risk. Sullivan et al. (1999) used the bootstrap methodology for the quantification of the biases in technical trading rules resulting from using information from the data to guide subsequent analysis of the same data (data-snooping biases). But in view of the anomalies mentioned earlier it is hard to justify the use of the stationary bootstrap of Politis and Romano (1994), which re-samples blocks of varying length from the original data.

More recent trading strategies are based on more sophisticated statistical methods. For example, Gencay (1998) and Qi (1999) presented evidence in favor of nonlinear predictability of stock returns. Both authors used neural network models and compared the performance of their models with conventional linear models. The same regressors were used for the linear and the nonlinear models. Since Qi’s set of explanatory variables does not only include financial variables but also economic variables like industrial output and inflation, there is the problem that the revised data used by the author were not available at the time when the investors had to make their investment decisions. In contrast, Gencay used only past prices and technical indicators (i.e., differences between prices and mov-

ing averages of the  $m$  most recent prices for  $m = 50, 200$ ). Neither author tried hard to justify his choice of the set of regressors. Qi used just the same set of nine regressors, which had already been used by Pesaran and Timmerman (1995) in a linear analysis of the same data set, and Gencay gave no reasons for his choice of the lengths of the moving averages. Both Gencay (1998) and Qi (1999) reported that the out-of-sample predictive performance of the neural network models was better than that of the linear models. But this does not necessarily mean that the neural network models are indeed superior. A comparison between different model classes with respect to their out-of-sample predictive performance can still be severely biased even if both estimation and identification are carried out strictly out of sample. An additional indispensable requirement is that only one identification method is considered. For example, trying out both AIC (Akaike's information criterion, Akaike, 1973) and BIC (Bayesian information criterion, Schwarz, 1978) for the determination of the dimension of the linear model and reporting only the results obtained with one of these criteria would clearly be unacceptable. Not surprisingly, things get even worse in case of nonlinear models. As is well known, the performance of neural networks depends critically on a number of subjective choices, e.g., the choice of the architecture (layers, units, transfer functions), the choice of the training algorithm, and the choice of the initial weights. It is therefore extremely difficult to assess the evidential value of the results obtained with these sophisticated models.

In view of the above discussion it is no wonder that the various claims of having found strong evidence against market efficiency are perceived by many researchers as greatly exaggerated (see, e.g., Malkiel, 1999). Obviously, there is still a need for more conclusive evidence. In Section 3, we consider possibly profitable trading strategies. These strategies are kept as simple as possible because they only serve as counter-examples to (the weak form of) the efficient market hypothesis. They are just rules for switching between a single risky asset and a risk-free asset. The extra profit is therefore not obtained at the price of bearing a higher risk. The trading decisions are solely based on the past history of the risky asset itself.

## 2. Higher-order Anomalies

### 2.1 Multiple nonstationarities

A useful method to detect nonstationarities in a time series is to calculate various statistics locally and to plot them cumulatively. For example, in order to check the stationarity of the returns

$$r_t = \log(DJI_t) - \log(DJI_{t-1}), \quad t = 1, 2, \dots, n$$

on the daily Dow Jones Industrial Average (DJI), we could first calculate the sample autocorrelation coefficient of order  $k$

$$\hat{\rho}_k[j] = \frac{\hat{\gamma}_k[j]}{\hat{\rho}_0[j]}, \quad (2.1)$$

where

$$\hat{\gamma}_k[j] = \frac{1}{q} \sum_{s=j}^{j+q-1-k} (r_s - \bar{r}[j])(r_{s+k} - \bar{r}[j]),$$

and

$$\bar{r}[j] = \frac{1}{q} \sum_{s=j}^{j+q-1} r_s,$$

for each of the  $n-q+1$  overlapping segments  $(r_j, \dots, r_{j+q-1})$ ,  $j = 1, 2, \dots, n-q+1$  of length  $q$  and then plot the sums

$$S_t = \frac{1}{n-q+1} \sum_{j=1}^t \hat{\rho}_k[j].$$

against time  $t$ . For  $k = 1, 2, \dots, 5$  and  $q = 250$  and we obtain the five graphs displayed in Figure 1. The data were downloaded from Yahoo! Finance. The overall first-order autocorrelation coefficient is close to 10%. Obviously, the magnitude of this coefficient changes over time. It was much smaller in the 1980s and 1990s than before. We should therefore calculate and interpret this coefficient only for seemingly stationary sub-periods in which there is a fairly linear increase in the cumulative first-order autocorrelation.

Methods for the division of a financial time series into stationary segments need not necessarily be based only on the first-order autocorrelation coefficient. Higher-order autocorrelation coefficients are also of interest. But this does not mean that we have to examine each autocorrelation coefficient separately, we could as well try to find instabilities in the spectral distribution function. Similarly, we could look for instabilities in the unconditional distribution rather than investigate individual numerical measures like the mean, the variance, the skewness, and the kurtosis. Many procedures for the detection of structural breaks are available (see, e.g., Andrews, 1993; Bai, 1994, 1996; Chu, 1995; Chu *et al.*, 1996; Hidalgo, 1995; Picard, 1985; Sowell, 1996) but they often are not robust against heavy tails or require that the locations of the possible changes are specified a priori or do not allow for dependence in the data. Recently, Inoue (2001) proposed a nonparametric test for distributional stability, which does not have any of these disadvantages. His test is based on the distance (Kolmogorov-Smirnov or

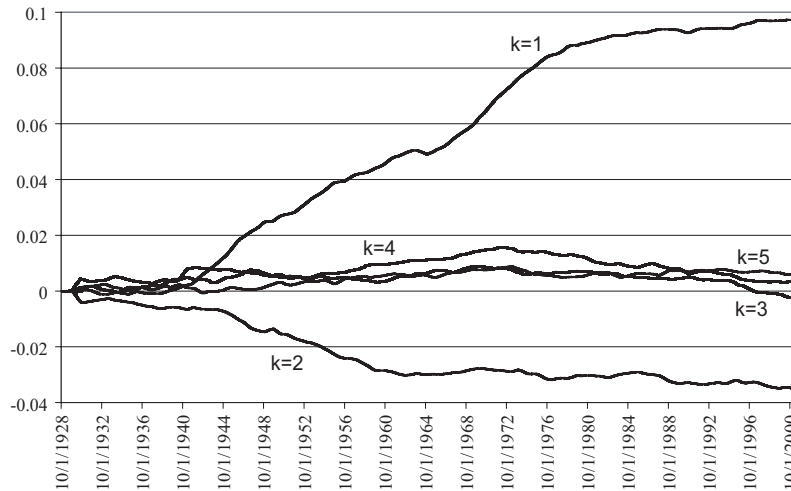


Figure 1: Cumulative local autocorrelations of order  $k = 1, \dots, 5$  calculated from segments of 250 returns on the DJI

Cramer-Von Mises) between the pre-change distribution function and the post-change distribution function. Applying his test to the weekly log-differences of several exchange rates and stock market indices (including the DJI from 1969 to 1996) he found evidence of distributional instability. If this instability is caused by a single structural break, the time point maximizing the Kolmogorov-Smirnov distance could serve as an estimate of the break location. In the more general case of multiple breaks, no meaningful estimates of break locations are provided. In addition, it may be expected that tests based on the Kolmogorov-Smirnov distance will have extremely low power in case of multiple breaks (see, e.g., Reschenhofer and Bomze, 1991). Both shortcomings could be rectified at the same time by replacing the Kolmogorov-Smirnov distance with the generalized Kolmogorov-Smirnov distance (Reschenhofer, 1997). But even if this could be accomplished somehow, we would still have to determine the number of breaks as well as to distinguish between structural breaks and other nonstationarities like smooth transitions.

Anyhow, when we have a closer look at Figure 1, we might even call the whole idea of stationary sub-periods into question. The second largest autocorrelation coefficient is that of order two. The size of this coefficient decreased already in the early sixties. There is no way that the changes in the first two autocorrelation coefficients could have occurred at the same time. It is to be feared that the number of possible breakpoints will increase further if we examine additional

features of this time series. On top of that it will be shown in the next subsection that covariance stationarity is even violated in very short periods of just a few days.

## 2.2 Calendar effects in higher-order moments

Structural breaks are not the only undesirable features of financial time series. It is well known that series of daily stock returns exhibit significant weekly and seasonal patterns. To examine whether the days of the week differ also with respect to higher-order characteristics we study the autocorrelation in the returns as well as in the squared returns separately for each day of the week. For example, to estimate the local autocorrelation of order one for Tuesday returns we consider the last 50 pairs  $(x_t, y_t)$  of (non-missing) Monday returns  $x_t$  and (non-missing) Tuesday returns  $y_t$  and calculate the local sample correlation coefficient

$$\hat{\rho} = \frac{\sum(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum(x_t - \bar{x})^2 \sum(y_t - \bar{y})^2}}$$

from this sample. All local sample correlation coefficients are then divided by the total number of local sample correlation coefficients and plotted cumulatively against time. Analogously we proceed for the other days of the week and also for the squared returns. Figure 2 shows that in the 1980s and 1990s the Thursday returns have been positively correlated with the Friday returns while the Monday returns have been negatively correlated with the Tuesday returns. Similarly, Figure 3 shows that in the 1970s, 1980s, and 1990s the squared Monday returns have been positively correlated with the squared Wednesday returns while the squared Tuesday returns have basically been uncorrelated with the squared Thursday returns. These findings suggest that the common approach of calculating autocorrelation coefficients, carrying out tests based upon autocorrelation coefficients, and estimating ARCH/GARCH models (Engle, 1982, Bollerslev, 1986) or stochastic volatility models (Taylor, 1986) is untenable in the analysis of daily financial data.

## 2.3 Concluding remarks

The evidence of multiple nonstationarities and higher-order day-of-the-week effects provided by Figures 1-3 is beyond doubt. Although carefully produced graphs can provide more conclusive evidence than significance tests that are based on implausible assumptions, some researchers are seriously addicted to  $p$ -values and keep on testing even in desperate situations, typically by resorting to the bootstrap methodology as a universal remedy. However, in the light of Figures 1-3 it seems unlikely that the stationary bootstrap of Politis and Romano (1994) is

appropriate for series of daily stock returns. It is also very questionable whether it is possible to modify conventional heteroskedasticity robust tests in order to make them insensitive to these kinds of anomalies. At least, we would have to add additional assumptions to the already long list of implausible and/or unverifiable assumptions required by these tests. As pointed out by Heckman (2001) there is a very real danger that such tests might be perceived by many applied economists as irrelevant to empirical research.

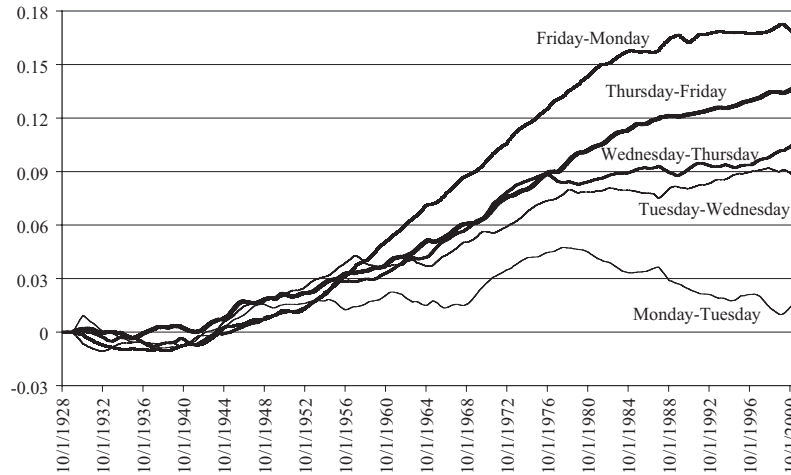


Figure 2: Cumulative local correlation coefficients (calculated from 50 pairs of returns) between Monday returns and Tuesday returns, Tuesday returns and Wednesday returns, etc.

### 3. Examining the Efficient Market Hypothesis

#### 3.1 Simple trading rules

To avoid controversial risk evaluations we consider only trading strategies that switch from the DJI to cash and vice versa. Using the DJI as the risky asset has the advantage that we can be quite sure that all stocks contained in this index are traded very frequently, particularly also near the open and near the close of the trading day. This reduces the risk that artificial autocorrelation is induced when some stocks are not traded towards the close and then tend to catch up with the other stocks on the next trading day (nonsynchronous trading effect). Anyhow, there is hardly any empirical evidence that the nonsynchronous trading effect is an important source of spurious autocorrelation (see Campbell et al., 1997). We



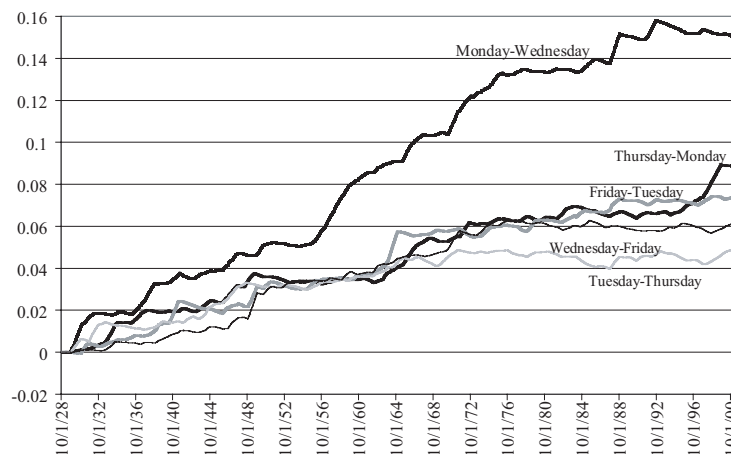


Figure 3: Cumulative local correlation coefficients (calculated from 50 pairs of squared returns) between squared Monday returns and squared Wednesday returns, squared Tuesday returns and squared Thursday returns, etc.

also do not need the nonsynchronous trading effect for the explanation of the fact that the strong autocorrelation observed in returns on the DJI is not present in returns on index futures. This discrepancy is simply due to nonstationarities in the DJI series and the fact that the trading with index futures started only in the 1980s. Note that the high autocorrelation in the returns on the DJI is basically due to the 1940s, 1950, 1960, and 1970s. In the 1980s and 1990s the autocorrelation was much smaller. The availability of long series of historical quotes is the main argument to use the DJI.

We will first consider strategies, which are profitable in the absence of transaction costs, and then try to achieve competitiveness in the harsh reality. At each stage, we will try to minimize the number of independently adjustable parameters in order to safeguard against data-snooping biases. Even with daily data we must still be aware of the risk of over-fitting. Using only a small number of parameters, e.g., one dummy variable for every time period in which the DJI suddenly lost a lot of value, we could easily construct strategies that would have been extremely profitable in the past, but are of no use to predict future crashes. We could use Alexander's (1961) filter rule as a starting point. A minor disadvantage of this rule is that its performance is highly sensitive to the choice of the threshold value  $v$ . While the use of a single parameter for the fine-tuning of a successful strategy for short-term trading is relatively harmless, the situation can quickly deteriorate, when we introduce additional parameters to improve its

performance. For that reason, we prefer to start with even simpler strategies, which do not depend on any parameter:

**Strategy 1:** Buy at the end of a good day ( $r_t > 0$ ) and sell at the end of a bad day ( $r_t < 0$ ).

**Strategy 2:** Buy at the end of a day that is better than the previous day ( $r_t > r_{t-1}$ ) and sell at the end of a day that is worse than the previous day ( $r_t < r_{t-1}$ ).

To keep things simple we assume that there are no partial transactions. We can either sell all we have or buy with all available cash. The starting capital is just the value of the DJI on the first day of the observation period. Strictly speaking, we do not face the possibility of ruin because all changes are in percentages. Nevertheless, a bad strategy can bring the capital very quickly close to zero. To implement Strategies 1 and 2 in practice, we must place suitable limit orders just before the end of the trading session. Alternatively, we could take a price stated shortly before the end of the trading session instead of the not yet available closing price as a basis for our investment decision. In the absence of transaction costs, both strategies are extremely successful (see Figure 4). Only at times when the DJI skyrocketed, the two strategies could not outperform the DJI. However, these facts are only convincing evidence in favor of predictability but not against market efficiency.

To prove market inefficiency we need to introduce transactions costs. Clearly, if we intend to follow our own trading strategy, we do not need the investment advice of a full service broker. Hence we can deal with a discount broker to minimize costs. If we trade on a large scale, we will have to pay less than 0.1% of the total market value of the transaction. Transactions costs for small investors have been much higher in the past, but it is hard to estimate the real costs for institutional investors throughout the century. In the following we assume that the one-way transaction costs are 0.1%. Even with these low costs Strategies 1 and 2 would have been ruinous. They simply trade too frequently. What we need is a choosier strategy. For additional hints about future price movements we might look at the four popular summary measures (open, high, low, close) of the intraday data. In case of a clear upward trend, we could expect that the opening price is close to the lowest price and the closing price is close to the highest price. Analogously, in case of a clear downward trend, we could expect that the opening price is close to the highest price and the closing price is close to the lowest price. In both cases, the absolute value of the quotient

$$q = \frac{\text{close} - \text{open}}{\text{high} - \text{low}}$$

should be large. To distinguish between large and small values we need a threshold value  $q^*$ . Instead of trying to find the “optimal” threshold value by maximizing the profit in the past, we prefer to use simply the sample mean  $|\bar{q}| = 0.35$  of the absolute values to avoid any suspicion of data-snooping.

**Strategy 3a:** Buy, if  $q_t > |\bar{q}|$ , and sell, if  $q_t < -|\bar{q}|$ .

Amazingly, it turns out that this primitive strategy is already quite competitive. Its performance is striking (see the upper half Figure 5). Overall, it outperforms the DJI. Relative to the DJI, it gains more in bear-markets than it loses in bull-markets. Its excellent performance from 1966 to 1976 casts serious doubts on market efficiency. Of course, arbitrageurs as well as nonstationarities will prevent any trading strategy from remaining profitable over a long period of time. For example, the performance of Strategy 3a may be affected by the fact that the distribution of  $|q_t|$  changes over the years. Its mean exhibits a clear downward trend, which may be attributed to the steady increase in the number of transactions per trading session. Besides the autocorrelation in the 1980s and 1990s was less significant than before.

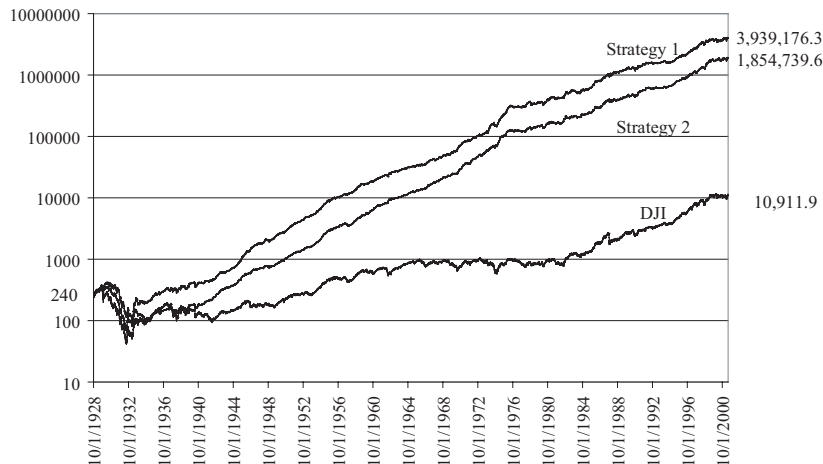


Figure 4: Performance of strategies 1 and 2 in the absence of transaction costs

We may take steps to improve the overall performance of Strategy 3a. First the threshold value could be increased to reduce trading frequency. Also two separate threshold values for buying and selling could be used. In this case, the threshold value for selling,  $q_-$ , should be greater than that for buying,  $q_+$ , to improve the poor performance in bull markets:

**Strategy 3b:** Buy, if  $q_t > q_+$ , and sell, if  $q_t < -q_-$ , where  $q_- > q_+ > |\bar{q}|$ .

For  $q_+ = 0.4, 0.5$  and  $q_- = q_+ + 0.1, q_+ + 0.2$  we obtain the four graphs displayed in the lower half of Figure 5. The overall performance has increased dramatically (the final wealth ranges from 25,559.2 to 41,195.5) but we should not pursue the fine-tuning too far. We might obtain over-optimistic results. For example, if we end up with very large values for both threshold values, the number of trades will be small and the chance of earning money just by luckily avoiding one or two of the big crashes will be high. Consequently, even a high profit would be of little evidential value.

In contrast, if our main interest is to earn money rather than to collect evidence against the efficient market hypothesis, we cannot afford being purists. In this case, we should from the beginning look at all the available information (e.g., trading volumes, dividends, earnings-price ratios, prices of other assets, interest rates, exchange rates, ...). Of course, that means arduous and detailed work. To make our life easier, we might be tempted to pass the buck to powerful nonparametric regression techniques (e.g., artificial neural networks) and automatic model selection criteria. But this is not recommended for anyone but a very experienced time-series modeler. The risk of finding spurious relationships is enormous. This is even true for the univariate case. Just imagine a sequence of plots of cumulative sample autocorrelation coefficients. As the lag-order increases the magnitudes of the coefficients will decrease and soon the plots will look like plots of random walks. Almost inevitably, some linear combination of these “random walks” will be very similar to the original price series. How can a stupid regression technique ever recognize that this similarity is coincidental.

### 3.2 Concluding remarks

Using a very simple and basically purely qualitative trading strategy (Strategy 3a) we managed to outperform (or even outclass) the buy-and-hold strategy for a long period of time (from 1966 to 1976). The special value of this empirical evidence against the efficient markets hypothesis is due to the simplicity of the strategy, the high trading frequency, and the length of the time period in which the strategy is highly profitable. In contrast, previous investigations of market efficiency mostly used trading strategies with several independently adjustable parameters and/or monthly data. However, even our study contains some fuzzy elements (e.g., our ignorance about the real transactions costs for institutional investors and our disregard for possibly non-negligible differences between the size of dividends and the size of short-term interest rates) that allow passionate advocates of market efficiency to play down the significance of our results.

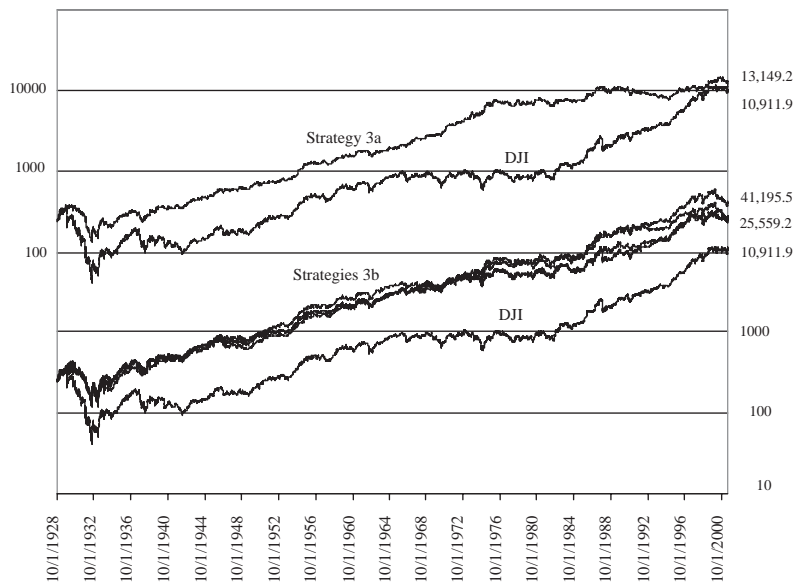


Figure 5: Performance of strategy 3a and four examples of strategy 3b with transaction costs of 0.1%

## References

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory* (Edited by B. N. Petrov and F. Csaki). Akademia Kiado, 267-281.
- Alexander, S. (1961). Price movements in speculative markets: Trends or random walks. *Industrial Management Review* **2**, 7-26.
- D. W. K. Andrews (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica* **61**, 821-856.
- J. Bai (1994). Weak convergence of the sequential empirical processes of residuals in ARMA models. *Annals of Statistics* **22**, 2051-2061.
- J. Bai (1996). Testing for parameter constancy in linear regressions: An empirical distribution function approach. *Econometrica* **64**, 597-622.
- T. Bollerslev (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* **31**, 307-327.
- Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.

- Chu, C.-S. J. (1995). Detecting parameter shift in GARCH models. *Econometric Reviews* **14**, 241-266.
- Chu, C.-S. J., Stinchcombe, M., and White, H. (1996). Monitoring structural change. *Econometrica* **64**, 1045-1065.
- Deo, R. S. (2000). Spectral tests of the martingale hypothesis under conditional heteroscedasticity. *Journal of Econometrics* **99**, 291-315.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987-1007.
- Fama, E. and Blume, M. (1966). Filter rules and stock market trading profits. *Journal of Business* **39**, 226-241.
- Gencay, R. (1998). The predictability of security returns with simple technical trading. *Journal of Empirical Finance* **5**, 347-359.
- Heckman, J. J. (2001). Econometrics and empirical economics. *Journal of Econometrics*, **100** 3-5.
- Hidalgo, J. (1995). A nonparametric conditional moment test for structural stability. *Econometric Theory* **11**, 671-698.
- Inoue, A. (2001). Testing for distributional change in time series. *Econometric Theory* **17**, 156-187.
- Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance* **45**, 881-898.
- Jegadeesh, N. and Titman S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* **48**, 65-91.
- Jensen, M. and Bennington, G. (1970). Random walk and technical theories: Some additional evidence. *Journal of Finance* **25**, 469-482.
- Lehmann, B. (1990). Fads, martingales and market efficiency. *Quarterly Journal of Economics* **105**, 1-28.
- Leuthold, R. M. (1972). Random walks and price trends: The live cattle futures markets. *Journal of Finance* **27**, 879-889.
- Levy, R. (1967). Relative strength as a criterion for investment selection. *Journal of Finance* **22**, 595-610.
- Lo, A. W. and MacKinlay, A. C. (1988). Stock prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies* **1**, 41-66.
- Malkiel, B. G. (1999). *A Random Walk Down Wall Street*, 2nd ed. W. W. Norton & Co.
- Pesaran, M. H. and Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *Journal of Finance* **50**, 1201-1228.

- 
- D. Picard (1985). Testing and estimating change-points in time series. *Advances in Applied Probability* **17**, 841-867.
- Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association* **89**, 1303-1313.
- Qi, M. (1999). Nonlinear predictability of stock returns using financial and economic variables. *Journal of Business and Economic Statistics* **17**, 419-429.
- Reschenhofer, E. (1997). Generalization of the Kolmogorov-Smirnov test. *Computational Statistics and Data Analysis* **22**, 433-441.
- Reschenhofer, E. and Bomze, I. M. (1991). Length tests for goodness-of-fit. *Biometrika* **78**, 207-216.
- Schwarz, G. (1978), Estimating the dimension of a model. *Annals of Statistics* **6**, 461-464.
- Stevenson, R. A. and Bear, R. M. (1970). Commodity futures: Trends or random walks? *Journal of Finance* **25**, 65-81.
- Sowell, F. (1996). Optimal tests for parameter instability in the generalized method of moments framework. *Econometrica* **64**, 1085-1107.
- Sullivan, R., Timmermann, A., and White, H. (1999). Data-snooping, technical trading rule performance and the bootstrap. *Journal of Finance* **54**, 1647-1692.
- Taylor, S. J. (1984). Estimating the variance of autocorrelations calculated from financial time series. *Applied Statistics* **33**, 300-308.
- S. J. Taylor (1986). *Modelling Financial Time Series*. John Wiley and Sons.

Received November 26, 2002; accepted March 11, 2003

Erhard Reschenhofer  
Department of Statistics and Decision Support Systems  
University of Vienna  
Universitätsstr. 5, A-1210 Vienna  
Austria  
erhard.reschenhofer@univie.ac.at