Predictable Patterns in Stock Returns

Thomas Hellström[1] and Kenneth Holmström[2]
Center of Mathematical Modeling (CMM)
Department of Mathematics and Physics
Mälardalen University
P.O.Box 883
S-721 23 Västerås, Sweden

Published as Opuscula ISRN HEV-BIB-OP-30-SE[3]

August 12, 1998

Abstract

This paper presents statistical investigations regarding the predictability of stock returns. The examined data covers 207 stocks on the Swedish stock market for the time period 1987-1996. The results show trend behavior and autocorrelation values that are stable even when the entire time interval is broken down to yearly intervals. A proposed concept of daily returns R_D and a comparision to the ordinary step returns R_S, show a significant difference in the data, caused by the holiday effect. It is also shown that seasonal variables, such as the month of the year, affect the stock returns more than the average daily changes. This is consequential for all methods where the seasonal variables are not taken into account when predicting daily stock returns.

Keywords: Data Mining, Finance, Prediction, Statistics, Stock returns.

[1]also Umeå University Umeå Sweden. email: thomash@cs.umu.se
[2]hkh@mdh.se
[3]Also available as Technical Report Series IMa-TOM-1997-09

1 Introduction

This paper presents results from a statistical analysis of stocks from the Swedish stock market for the period 1987-1996. The purpose of the work is to examine, in general terms, the statistical properties that are relevant to the development of prediction algorithms. Therefore, at this stage, we are not aiming at extracting patterns or dependencies that can be utilized directly for trading purposes.

The following properties of the price generating process will be considered in the analysis:

- The process behaves very much like a random walk process. The autocorrelation for day to day changes is very low. Hawawini and Keim [3] conclude on several studies that the serial correlation in stock price time series is economically and statistically insignificant.
- The process is 'regime shifting', in the sense that the underlying process varies over time. The noise level and volatility in the time series change as the stock markets move in and out of periods of "turbulence", "hausse", and "baisse". This causes great problems for traditional algorithms of time series predictions.

The patterns that may be found in the data are most surely very weak and the risk of "data snooping", i.e. interpreting random patterns as properties of the process, is extremely high. The statistical analysis takes that into consideration, even if the overall question of generalization should be left to the prediction task, which is not covered in this paper.

1.1 Common viewpoints

The efficient market hypothesis (EMH) states that the current market price reflects the assimilation of all the available information. A direct consequence of the strongest form of EMH is that the prices follow a random walk pattern and that future prices can not be predicted. The hypothesis has been supported in a number of research studies, see e.g. [8],[5], and has for long been the "official" viewpoint among many academicians. However, in recent years, there have been many published papers showing that the efficient market hypothesis is far from correct. Fama [2] declares, in a review of the market efficiency literature, that the efficient market hypothesis is most surely false.

Many of the market actors claim that they can predict the stock prices in a way that would make them a profit. Almost all professional traders rely to some extent on either technical or fundamental analysis of past information. They buy when the market is 'bullish' and sell when it is 'bearish'; thereby assuming a direct correlation between the current trend and future prices. There are also numerous research papers claiming that applying non linear models such as neural networks and genetic algorithms *can* be used for successful predictions.

2 Descriptive analysis of data

The data examined in this paper is picked from the Swedish stock market for the period 1987-1996. Results for two sets of stocks are normally presented; *SXG* which is a compilation of 32 major stocks with active trading and *SXBIG* which is a compilation of 207 major and minor stocks (including those in *SXG*). The stocks included in *SXG* are listed in Table 1.

The k-step return $R_k(t)$ of a stock price time series $y(t)$ is defined as

$$R_k(t) = 100 \cdot \frac{y(t) - y(t-k)}{y(t-k)}. \tag{1}$$

The returns $R_k(t)$ are the primary target in most research on the predictability of stocks. Some reasons are:

1. $R_k(t)$ has a relatively constant range even if many years data has been used as input. The prices $y(t)$ obviously vary a lot more and make it difficult to create a valid model for a longer period of time.
2. $R_k(t)$ for different stocks may also be compared on an equal basis (this is however seldom done in published research).
3. It is easy to evaluate a prediction algorithm for $R_k(t)$ by computing the prediction accuracy of the sign of $R_k(t)$. A long time accuracy above 50% (or more precisely above the historical mean) indicates that a true prediction has taken place.

The basic statistical properties of $R_k(t)$ for the two sets of stocks are listed in tables 2 and 3. The values in the tables are mean values for the included stocks. Each column shows data for one particular value of k.

The last six lines in the tables show the distribution of signs for the returns. "$Return = 0$" is the fraction of returns that are equal to zero. "$Return > 0$" is the fraction of returns that are greater than zero and "$Return < 0$" is the fraction of returns that are less than zero. *"Up fraction"* is computed as

$$100 \cdot \frac{\text{"}Return > 0\text{"}}{\text{"}Return > 0\text{"} + \text{"}Return < 0\text{"}} \quad , \tag{2}$$

which is the positive fraction of all non-zero moves. *"Up fraction"* is a relevant measure, when it comes to evaluating the hit rate of prediction algorithms. Looking at one-step returns in the tables, the "*Up fraction*" for the *SXG* stocks is 50.8% and for the *SXBIG* stocks is 50.6%. The "*Mean Up*" and "*Mean Down*" columns shows the mean value of the positive and negative returns respectively.

The fractions of zero returns in the data material are somewhat surprisingly high, 14.1% for the *SXG* stocks and 23.4% for the *SXBIG* stocks. The higher value in the latter set is related to the lower degree of activity in the smaller stocks included in *SXBIG*. The

Table 1: Stocks in set SXG

	From	To
AGA A	870102	961230
ABB A	870102	961230
AssiDoman	940408	961230
Astra A	870102	961230
Atlas Copco A	870102	961230
Autoliv	940609	961230
Avesta Sheffield	870126	961230
Electrolux B	870102	961230
Ericsson B	870102	961230
Hennes & Mauritz	870102	961230
Incentive A	870102	961230
Industrivard A	870116	961230
Investor A	870102	961230
Kinnevik B	870105	961230
MoDo B	870102	961230
Pharm & Up SDB	871015	961230
SE-Banken A	870102	961230
Sandvik A	870102	961230
SCA B	870102	961230
SHB A	870102	961230
Skandia	870102	961230
Skanska B	870102	961230
SKF B	870102	961230
Sparbanken A	950609	961230
SSAB A	890703	961230
Stora A	870102	961230
Sydkraft C	870102	961230
Trelleborg B	870102	961230
Trygg-H SPP B	891208	961230
Volvo B	870102	961230
Allgon B	880527	961230
Nokia A	870102	961230

Table 2: Mean k-step returns for 32 major Swedish stocks (SXG)

	k						
	1	**2**	**5**	**10**	**20**	**50**	**100**
Mean	0.091	0.183	0.445	0.892	1.831	4.585	9.142
Median	0.000	0.017	0.140	0.494	1.447	4.025	7.133
Std. dev	2.18	3.17	5.09	7.29	10.65	18.23	28.42
Skewness	0.52	0.70	0.78	0.71	0.56	0.53	0.64
Kurtosis	12.77	12.36	12.46	10.86	8.61	6.12	5.51
No of points	2178	2175	2172	2165	2153	2124	2074
Returns=0 (%)	14.1	9.3	5.5	3.6	2.4	1.2	0.8
Returns¿0 (%)	43.6	46.1	49.5	52.1	56.1	60.2	62.6
Returns¡0 (%)	42.2	44.6	45.0	44.3	41.5	38.5	36.6
Up fraction (%)	50.8	50.9	52.3	54.0	57.5	61.0	63.1
Mean Up	1.8	2.6	4.1	5.9	8.6	15.3	24.8
Mean Down	-1.6	-2.2	-3.5	-4.8	-7.0	-11.5	-16.0

Table 3: Mean k-step returns for 207 Swedish stocks (SXBIG)

	k						
	1	**2**	**5**	**10**	**20**	**50**	**100**
Mean	0.143	0.274	0.585	1.058	2.007	4.584	8.651
Median	0.000	0.007	0.060	0.248	0.946	2.895	5.148
Std. dev	3.02	4.15	6.15	8.42	11.80	18.82	27.80
Skewness	0.79	1.06	1.02	0.93	0.83	0.78	0.82
Kurtosis	15.78	16.49	11.55	9.27	7.58	5.99	5.59
No of points	1367	1363	1356	1347	1333	1306	1259
Returns=0 (%)	23.4	17.0	10.8	7.5	4.9	2.7	1.9
Returns¿0 (%)	38.7	42.0	45.6	48.5	52.1	56.1	57.6
Returns¡0 (%)	37.9	41.1	43.6	44.1	43.0	41.2	40.5
Up fraction (%)	50.6	50.6	51.1	52.3	54.7	57.6	58.7
Mean Up	2.7	3.5	5.2	7.1	10.1	16.8	26.5
Mean Down	-2.3	-2.9	-4.0	-5.3	-7.3	-11.3	-15.5

zero returns must be dealt with in a proper way when evaluating hit rates for prediction algorithms. The "*Up fraction*" circumvents the zero returns by simply removing them before calculating the hit rate. In this way, the zero returns will be counted as both increases and decreases, in equal proportions. A similar procedure is proposed in [4] for a test metric when doing stock predictions.

3 Autocorrelation

Much of the literature about market efficiency and stock price predictability is focused on the properties of the autocorrelation function of the stock returns. Fama [2] surveys a number of investigations and concludes that "The new research produces precise evidence on the predictability of daily and weekly returns from past returns". In this section we present our own results from investigations of data from the Swedish stock market.

The autocorrelation for a time series X describes how well X is correlated to itself at two different times. For a stationary X the autocorrelation is defined as a function of the time difference τ as

$$A_X(\tau) \;=\; E[X(t)X(t+\tau)]. \tag{3}$$

The autocorrelation $A_X(\tau)$ can be viewed as a measure of the τ-step memory in the process generating X, computed as a mean value over the entire time series. By computing $A_X(\tau)$ for $\tau = 1, 2, 3, ...$, the autocorrelation can be plotted as a function of τ. It should be noticed that only linear dependencies are fully revealed by the autocorrelation. It is however often used as a general indicator of predictability, even for nonlinear functions.

We want to investigate the autocorrelation for the one-step stock returns $R(t)$ defined in (1). The autocorrelation for $R(t)$ (denoted ACF) is computed using (3).

Figure 1 shows the ACF for the *SXG* stocks. For each lag (value on τ), the mean value of the ACF for the 32 stocks is plotted. To test the stationary state assumption for the underlying process, the data is broken down to five equal blocks, each covering about two years of stock data. The graph to the left in Figure 1 shows the ACF for the entire time period. The graph to the right shows one curve for each two year period.

The autocorrelation values are overall very low! The autocorrelation for lag 1 is clearly positive, whereas the following lags, 2 and 3, are negative. The following lags seem to vary in a rather random fashion around zero.

The same type of diagram for the larger *SXBIG* set is shown in Figure 2. Surprisingly, the result for these 207 stocks shows a significantly *negative* correlation in the first lag, whereas *positive* correlation is found for the 32 major stocks in *SXG*. We do not give any explanation for this observation but it is reasonable to suspect a "small-company effect" in the larger *SXBIG* set.

As a reference, the corresponding graphs for 100 simulated random walk stocks are presented in Figure 3. The autocorrelation is very close to zero in a statistically stable way. This strengthens the results shown in the previous graphs for the real stock data.

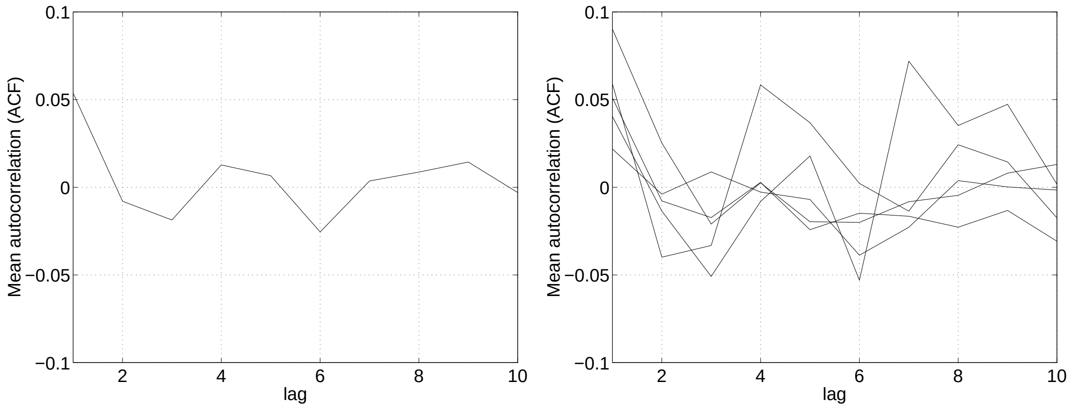

Figure 1: The autocorrelation for 32 major Swedish stocks (*SXG*) for the years 1987 to 1996. The right graph is broken down to five two-year curves

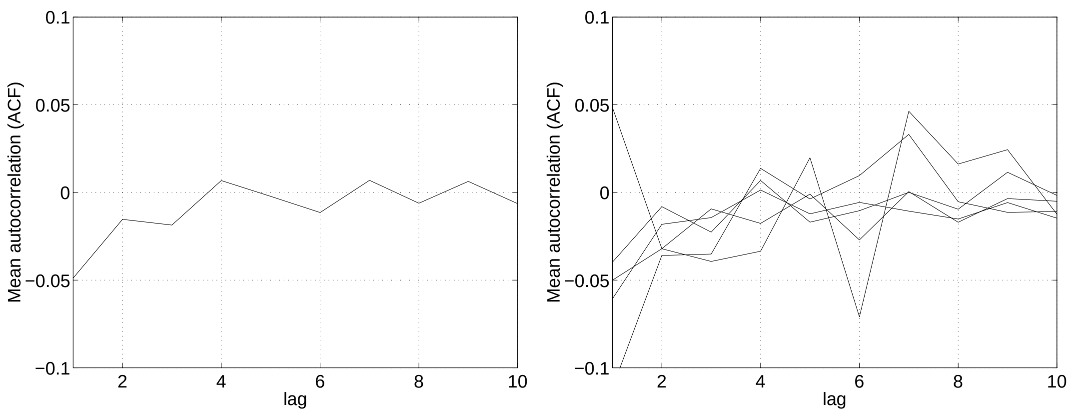

Figure 2: The autocorrelation for 207 Swedish stocks (*SXBIG*) for the years 1987 to 1996.

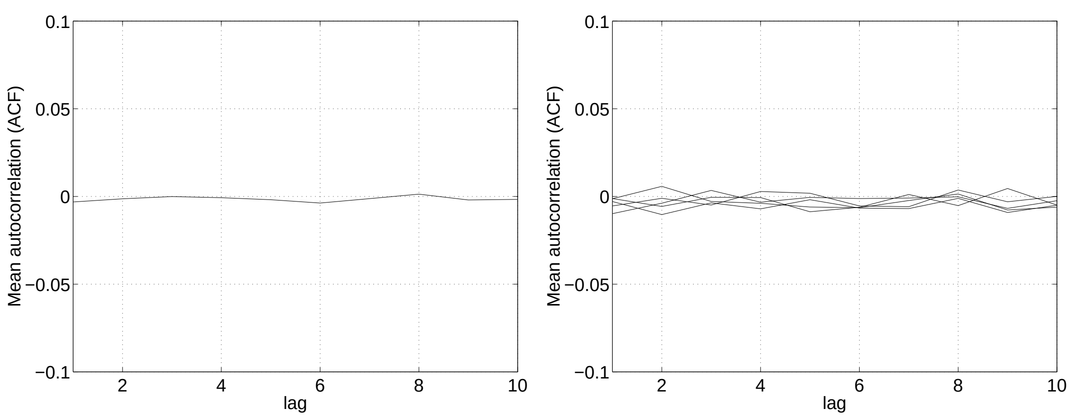

Figure 3: The autocorrelation for 100 simulated random walk stocks for the years 1987 to 1996

4 Trends

A trend following trading strategy means buying stocks, which have shown a positive trend for the last days, weeks or months. It also suggests selling stocks, which have shown a negative trend. In this section the relevance in such a strategy will be tested statistically.

A trend $T_k(t)$ is defined using the k-step return as

$$T_k(t) = \frac{100}{k} \cdot \frac{y(t) - y(t-k)}{y(t-k)}. \tag{4}$$

By setting different values on k we get measures telling how much the stock has increased per day since its value k days ago.

To see if $T_k(t)$ is connected to future changes, define the profit $P_h(t)$ computed h days ahead as

$$P_h(t) = 100 \cdot \frac{y(t+h) - y(t)}{y(t)}. \tag{5}$$

$P_h(t)$ is obviously equal to $R_h(t+h)$ (i.e. it is achieved by shifting the returns h days backwards).

In Table 4 the mean profit $P_1(t)$ is tabulated as a function of the trend $T_k(t)$, i.e. 1-step-forward profit versus k-step-back trends. Results for the *SXG* stocks for the years 1987-1996 are presented. Table 5 shows the "*Up fraction*" (2) and Table 6 the number of observations in each table entry. Tables 7, 8 and 9 show the same, but with $P_5(t)$ and $T_k(t)$, i.e. 5-step-forward profit per day versus k-step-backward trends per day.

The six tables 10, 11 12 13, 14 15 repeat the same analysis for the larger *SXBIG* set.

To ensure that found patterns reflects fundamental properties of the process generating the data, and not only idiosyncrasies in the data, the relation between trends and future returns are also presented in graphs with one curve representing one year. Figures 4 shows 1-step profits versus 1-step trends for *SXG* and *SXBIG* respectively. Figure 5 shows 2-step profit versus k-step trends. Figure 6 shows 5-step profit versus k-step trends.

Let us draw some conclusions from these statistical examinations of trends.

- The massive better part of returns falls into a region, where it is very difficult to claim any correlation between past and future price changes. The regions, where any correlation may be significant, are the sparsely populated extreme ones. Some interesting effects can however be observed. Looking at Table 11 we observe that a 5% decrease in price since yesterday, stands a 61.2% probability of showing an increase by tomorrow. The same negative correlation holds for both sets of investigated stocks.

- The cases with large increase since yesterday call for a more complex interpretation. Looking at figure 5, a difference between the two investigated sets of stocks becomes apparent. For the stocks in *SXG*, a 5% increase in price since yesterday, stands a 54.9% probability of showing an *increase* by tomorrow. For the stocks in *SXBIG*, a 5% increase in price since yesterday, stands a 53.3% probability of showing a *decrease* by tomorrow. This is in accordance with the observed difference in the first lag of the ACF, reported in Section 3.

5 Seasonal effects

In this section we present some empirical investigations showing how the day of the week and the month of the year affect the stock price returns. The results confirm and complement similar investigations on other stock markets world wide.

5.1 Handling the missing Weekends

The exchanges in most countries are closed on Saturday and Sunday. The price time series will therefore have gaps for these days if plotted against a calendar-time axis. The situation is normally approached in either of the two following ways:

- The returns from Friday to Monday are assumed to be three times the returns on each of the other days of the week. This is called the calendar-time hypothesis

- Returns are generated only during active trading days, and average returns are the same for all five days of the week. This is called the trading-time hypothesis.

Hawawini and Keim [3] report on negative Monday returns (i.e. relative price changes from Friday close to Monday close) for several European stock markets. This is inconsistent with both the calendar-time hypothesis and the trading-time hypothesis. A satisfactory explanation for this weekend effect does not, according to Hawawini and Leim, exist at present.

Our own results do not support any of the mentioned hypotheses either. However, the day-to-day returns exhibit a complex pattern, as shown in the next section. To proceed we need to define several metrics for a stock price time series $\{y(0), y(1), y(2), ..., y(N)\}$:

The **one-step** returns $R_S(t)$ are defined as the relative increase in price since the prior point in the time series, i.e.

Table 4: Mean 1-step returns for 32 stocks in SXG

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	0.52	-0.03	-0.05	-0.05	-0.12	-0.08	0.00	0.11	0.19	0.28	0.37	0.40	0.73
2	1.19	0.20	0.13	-0.09	-0.06	-0.08	0.03	0.13	0.24	0.31	0.28	0.31	0.43
3	1.87	0.06	0.15	0.16	-0.04	-0.07	0.03	0.15	0.20	0.25	0.27	-0.02	0.38
4	3.17	0.48	-0.03	0.12	-0.00	-0.04	0.03	0.16	0.20	0.16	0.42	-0.04	0.83
5	3.47	0.93	0.19	-0.03	0.02	-0.01	0.05	0.15	0.20	0.14	0.21	0.57	1.12
10	10.64	7.39	0.38	-0.11	0.00	-0.00	0.06	0.15	0.17	0.30	0.29	0.44	0.56
20			8.25	1.23	-0.14	-0.06	0.09	0.14	0.21	0.31	0.06	3.28	1.29
30				2.63	0.02	-0.07	0.09	0.13	0.22	0.33	2.19	1.24	1.49
50					0.48	-0.08	0.08	0.13	0.11	1.10	-0.17	1.95	1.21
100					4.68	0.18	0.05	0.13	0.26	0.57	0.78	0.29	0.81

Table 5: Fraction up/(up+down) moves (%) for stocks in SXG

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	55.9	50.6	49.6	48.5	47.1	47.3	49.1	51.4	53.3	54.9	55.6	55.0	54.9
2	58.3	51.9	53.1	49.5	47.6	47.8	49.5	51.8	54.5	54.4	52.9	51.8	48.8
3	62.2	51.4	49.5	53.3	48.7	47.3	49.8	52.6	53.6	53.6	51.5	42.4	47.9
4	64.8	55.7	48.9	50.4	49.9	47.9	50.0	52.8	54.1	50.8	49.4	43.0	50.6
5	60.6	58.9	51.0	49.8	50.0	48.9	50.1	52.8	52.7	50.0	48.3	49.1	53.8
10	100.0	73.3	52.6	50.3	49.3	48.8	50.8	52.8	50.4	51.8	50.2	47.5	43.5
20			77.8	56.1	47.1	47.5	51.4	52.3	50.7	48.6	49.1	50.0	55.6
30				56.1	47.6	47.3	51.5	51.7	50.2	50.7	55.8	61.1	54.8
50					52.3	46.3	51.2	51.8	48.8	54.4	52.8	57.1	50.0
100					71.4	49.5	50.5	51.6	49.2	54.8	64.7	50.0	53.5

Table 6: Number of points (SXG)

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	1396	1178	2539	6427	8701	7543	12373	7317	8508	6443	3048	1504	1807
2	492	561	1553	4644	9334	10822	12800	10219	9335	5399	2008	759	796
3	244	346	985	3514	8807	11837	15369	11419	9238	4575	1415	478	456
4	116	215	716	2717	8246	12485	17295	12122	9272	3865	950	317	310
5	71	136	553	2178	7478	13224	18504	12871	9185	3215	702	244	218
10	3	30	165	957	4794	13670	23902	15643	7177	1536	292	93	79
20	0	0	9	359	2548	12327	29970	17393	4441	742	123	18	27
30	0	0	0	70	1855	10829	33884	17278	3156	473	48	18	32
50	0	0	0	0	765	8802	39328	15867	1971	216	41	24	42
100	0	0	0	0	7	5798	46293	12148	861	258	40	5	48

Table 7: Mean 5-step returns for 32 stocks in SXG

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	1.14	0.05	0.44	0.18	0.21	0.14	0.31	0.43	0.65	0.67	0.64	0.74	1.22
2	1.87	0.80	0.67	0.29	0.23	0.23	0.26	0.50	0.62	0.68	0.44	0.61	1.36
3	2.66	0.68	0.88	0.62	0.30	0.15	0.29	0.52	0.60	0.65	0.60	0.30	1.57
4	4.28	0.89	1.45	0.52	0.27	0.17	0.35	0.52	0.66	0.50	1.15	-0.41	2.55
5	6.30	1.44	1.09	0.30	0.28	0.23	0.35	0.56	0.62	0.56	0.52	1.69	2.17
10	9.84	5.09	1.27	0.01	0.21	0.16	0.41	0.58	0.74	0.72	1.03	0.51	2.69
20			16.02	3.05	-0.27	-0.09	0.44	0.66	0.78	1.27	-1.05	2.80	8.47
30				6.04	-0.13	-0.21	0.47	0.60	0.86	1.67	6.70	-1.06	6.76
50					1.84	-0.26	0.39	0.59	0.76	3.10	4.13	9.78	4.26
100					7.96	1.01	0.25	0.51	1.32	3.25	1.91	4.36	3.81

Table 8: Fraction up/(up+down) moves (%) for stocks in SXG

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	56.7	51.3	51.2	50.7	50.4	50.0	52.0	52.7	54.6	54.1	53.4	52.5	51.4
2	55.9	55.4	55.3	51.7	50.3	51.1	51.5	53.0	54.0	54.1	50.6	50.2	49.3
3	56.4	54.5	55.6	55.0	51.5	50.1	51.9	53.2	53.7	53.2	49.8	48.7	46.1
4	62.6	55.3	59.1	54.8	51.8	50.5	52.1	52.8	54.2	50.7	51.6	43.4	50.2
5	65.2	55.5	57.1	53.3	52.0	51.0	52.0	53.4	53.5	50.7	48.0	52.2	47.5
10	66.7	62.1	50.3	49.3	51.4	49.7	52.9	53.8	52.7	51.1	50.4	51.2	41.3
20			100.0	59.2	46.6	48.1	53.4	54.3	50.3	50.5	40.2	50.0	56.0
30				69.1	46.1	47.4	53.5	53.0	51.7	52.2	63.0	41.2	62.1
50					55.7	45.7	52.8	53.4	49.1	60.0	63.4	63.6	64.1
100					83.3	51.0	51.7	52.9	54.5	64.8	61.1	80.0	63.6

Table 9: Number of points (SXG)

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	1382	1173	2549	6406	8682	7519	12367	7288	8467	6415	3022	1494	1793
2	486	558	1536	4638	9315	10783	12763	10172	9300	5369	2001	753	793
3	237	340	966	3507	8776	11792	15363	11392	9204	4541	1400	469	458
4	113	209	709	2710	8216	12459	17253	12102	9216	3832	934	315	312
5	71	133	547	2177	7442	13193	18487	12846	9143	3186	701	244	211
10	3	30	164	946	4774	13595	23856	15642	7154	1529	289	90	80
20	0	0	8	361	2533	12329	29881	17312	4434	735	123	17	27
30	0	0	0	70	1850	10819	33799	17193	3145	467	48	17	32
50	0	0	0	0	760	8777	39226	15803	1960	214	43	24	42
100	0	0	0	0	7	5771	46183	12106	861	259	40	5	48

Table 10: Mean 1-step returns for 207 stocks in SXBIG

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	1.42	0.40	0.30	0.09	-0.02	-0.07	0.02	0.09	0.11	0.14	0.15	0.16	0.06
2	2.35	0.75	0.44	0.22	0.06	-0.00	0.02	0.06	0.09	0.13	0.06	0.10	0.05
3	3.38	0.95	0.62	0.35	0.12	0.01	0.02	0.07	0.11	0.11	-0.00	-0.12	-0.12
4	4.47	1.36	0.65	0.45	0.14	0.04	0.02	0.07	0.12	0.04	-0.00	0.07	-0.03
5	5.06	2.21	0.87	0.40	0.20	0.04	0.04	0.10	0.10	0.05	-0.07	-0.03	-0.01
10	8.28	5.02	1.88	0.57	0.24	0.08	0.06	0.10	0.11	0.13	0.18	0.58	-0.89
20	40.22	11.58	2.43	1.84	0.26	0.05	0.08	0.11	0.16	0.15	-0.10	-0.02	-1.07
30			8.07	3.41	0.38	0.09	0.08	0.10	0.19	0.21	-0.12	0.59	-1.37
50				7.55	0.81	0.08	0.10	0.14	0.11	0.01	0.37	-0.71	-1.05
100					2.51	0.22	0.08	0.15	0.31	-0.15	-0.53	-2.10	-0.30

Table 11: Fraction up/(up+down) moves (%) for stocks in SXBIG

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	61.2	56.9	55.0	52.1	49.6	48.0	48.9	50.2	50.0	49.8	49.8	50.0	46.7
2	62.8	57.6	56.9	53.7	51.1	49.5	49.2	49.3	49.8	50.1	48.3	47.5	45.7
3	65.0	58.4	56.5	55.7	51.9	49.4	49.2	49.7	50.1	49.4	47.3	44.6	43.8
4	64.6	58.9	55.5	55.2	52.5	50.1	49.0	49.9	51.0	48.3	46.1	45.8	44.7
5	65.4	60.8	55.7	54.5	52.8	50.0	49.5	50.5	50.2	47.8	44.9	45.5	44.1
10	65.0	65.1	56.0	53.0	51.9	50.6	50.3	50.7	48.9	47.8	46.2	47.9	38.4
20	66.7	65.4	54.8	56.9	50.2	49.5	50.8	50.7	49.1	47.8	43.7	44.0	40.6
30			55.2	57.5	49.4	49.3	51.0	50.4	49.3	47.3	39.8	53.6	37.7
50				57.7	50.5	48.4	50.7	51.0	48.2	46.9	47.9	37.5	41.7
100					52.4	49.1	50.3	50.7	49.3	46.9	47.7	31.6	50.0

Table 12: Number of points (SXBIG)

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	10256	6255	12158	23991	29276	19134	60845	18447	28262	23213	13091	7077	13141
2	4071	3206	7475	19783	34386	35228	53625	33001	32519	20794	9214	4335	6703
3	2121	2101	5000	15661	33500	42333	58688	39244	32600	18332	6854	2994	4244
4	1187	1426	3640	12431	31945	46192	64660	42539	32581	15915	5432	2264	2982
5	744	990	2852	10142	29791	49017	69243	45101	32426	13989	4384	1698	2344
10	116	255	1047	4837	19894	52570	89411	53333	27421	8062	2238	807	890
20	3	28	168	1941	11475	47613	112230	59258	19290	4344	1029	338	320
30	0	0	32	708	8339	42626	127005	58767	14641	3021	625	209	179
50	0	0	0	59	4033	34769	147072	54372	10147	1655	311	115	83
100	0	0	0	0	364	23997	167490	44255	5972	1094	129	20	53

Table 13: Mean 5-step returns for 207 stocks in SXBIG

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	2.30	0.82	0.66	0.38	0.27	0.18	0.41	0.47	0.53	0.56	0.54	0.63	0.21
2	4.32	1.56	0.86	0.57	0.32	0.26	0.40	0.49	0.55	0.48	0.27	0.56	-0.16
3	5.89	1.67	1.36	0.70	0.38	0.24	0.41	0.51	0.54	0.50	0.26	0.61	-0.22
4	7.51	3.37	1.42	0.82	0.37	0.26	0.41	0.51	0.60	0.41	0.47	0.36	-0.16
5	9.31	3.75	1.95	0.73	0.39	0.28	0.40	0.60	0.55	0.46	0.30	0.63	-0.56
10	16.56	12.35	4.29	0.86	0.33	0.22	0.47	0.60	0.65	0.44	0.84	0.47	-3.16
20	78.89	11.97	13.65	4.96	0.18	0.06	0.44	0.63	0.76	1.36	0.20	-0.42	-4.85
30			23.96	8.33	0.64	0.08	0.45	0.59	0.95	1.27	1.61	-2.39	-6.15
50				12.18	2.39	0.01	0.41	0.73	0.71	0.92	0.73	-2.77	-5.85
100					8.33	0.68	0.33	0.68	1.08	0.58	-1.84	-4.26	2.10

Table 14: Fraction up/(up+down) moves (%) for stocks in SXBIG

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	58.6	54.4	53.8	51.3	50.4	49.6	50.2	52.3	51.4	50.7	49.4	48.7	44.1
2	61.6	56.3	55.9	53.4	50.7	50.1	50.3	51.3	51.2	49.7	46.5	46.8	41.5
3	63.5	57.1	56.9	54.9	51.6	49.9	50.6	51.2	50.8	48.6	45.6	44.4	41.3
4	63.5	60.7	56.4	55.5	51.9	50.2	50.6	51.1	51.2	47.4	45.1	43.1	40.7
5	64.7	62.5	56.5	54.7	52.4	50.1	50.7	51.7	50.4	46.9	44.5	44.1	39.2
10	74.0	67.5	57.4	52.1	51.3	49.8	51.4	51.8	49.6	45.9	44.5	43.7	34.6
20	100.0	53.6	66.9	58.6	48.3	48.5	51.6	51.9	49.2	48.9	40.4	38.4	31.6
30			58.1	60.2	48.0	48.3	51.8	51.0	49.4	47.7	41.6	35.7	30.3
50				47.4	51.4	46.8	51.2	52.3	47.3	45.9	43.6	38.5	38.7
100					58.2	48.8	50.4	51.9	49.7	45.3	44.5	36.8	58.3

Table 15: Number of points (SXBIG)

	k-day trend (%/day)												
k	**-5.00**	**-4.00**	**-3.00**	**-2.00**	**-1.00**	**-0.50**	**0.00**	**0.50**	**1.00**	**2.00**	**3.00**	**4.00**	**5.00**
1	10120	6203	12085	23788	29099	19057	60477	18284	27958	22945	12888	6969	12834
2	4000	3177	7442	19726	34260	35003	53355	32659	32188	20508	9050	4255	6453
3	2029	2081	4916	15631	33442	42073	58477	38851	32278	17965	6713	2932	4049
4	1137	1391	3576	12334	31909	45973	64282	42165	32168	15663	5278	2197	2862
5	713	965	2790	10055	29701	48998	68890	44674	31987	13759	4290	1648	2234
10	111	256	1010	4736	19775	52317	88853	53000	27101	7880	2200	782	862
20	2	28	167	1919	11375	47582	111654	58661	19034	4241	999	327	305
30	0	0	31	693	8232	42555	126516	58131	14367	2957	602	207	175
50	0	0	0	59	3971	34558	146319	53863	9990	1623	311	111	83
100	0	0	0	0	359	23859	166617	43693	5919	1079	121	20	52

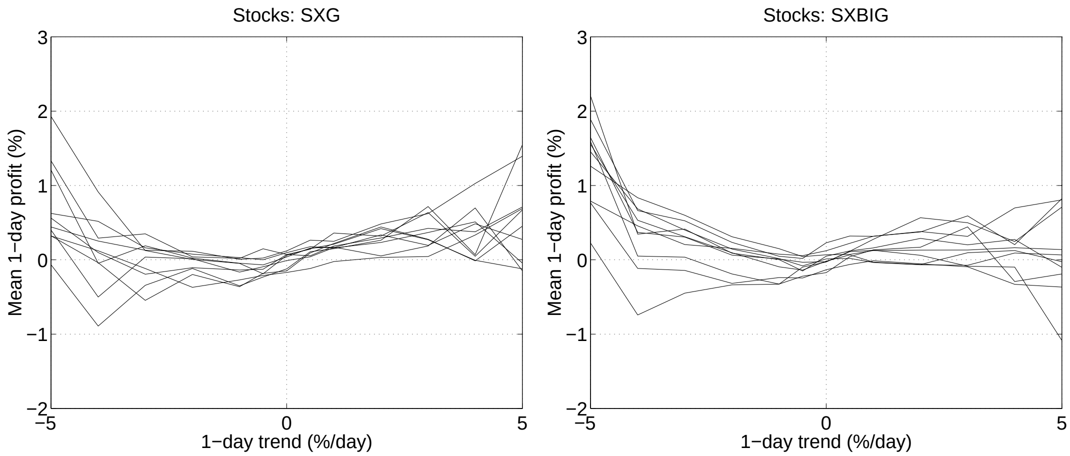

Figure 4: 1-step profits versus 1-step returns for stocks *SXG* and *SXBIG*. Each curve represents one year between 1987 and 1996.

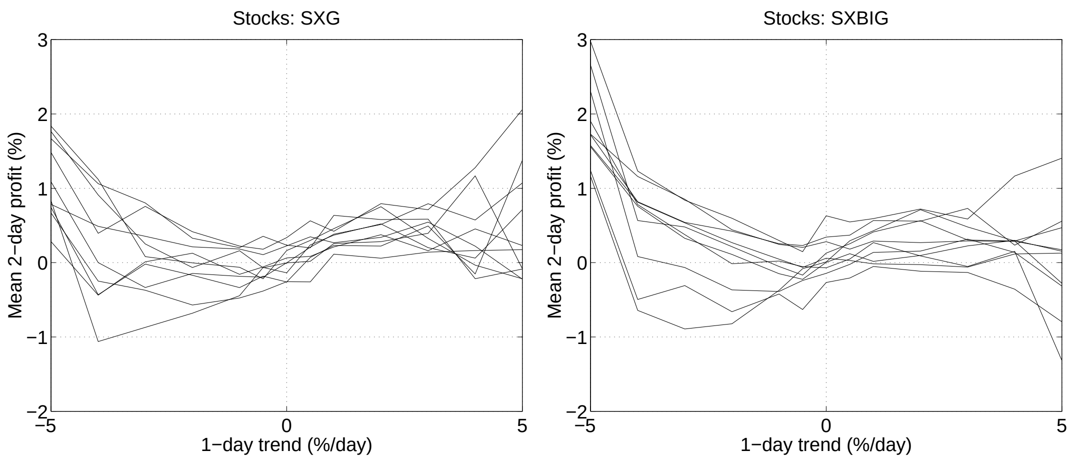

Figure 5: 2-step profits versus 1-step returns for stocks *SXG* and *SXBIG*. Each curve represents one year between 1987 and 1996.

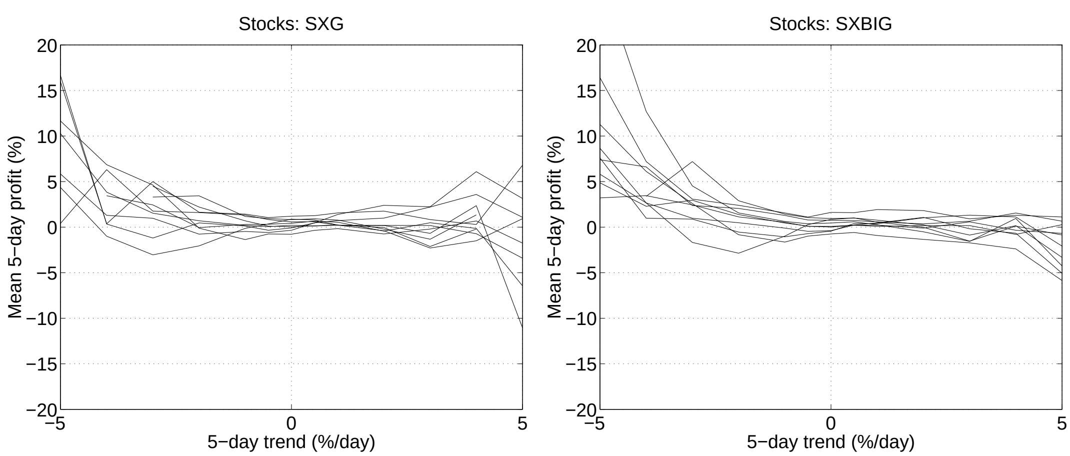

Figure 6: 5-step profits versus 5-step returns for the *SXG* and *SXBIG* stocks. Each curve represents one year between 1987 and 1996.

$$R_S(t) = 100 \cdot \frac{y(t) - y(t-1)}{y(t-1)}. \tag{6}$$

The mean of **one-step** returns $R_S(t)$ is computed using the standard definition as

$$R_S = \frac{1}{N}\sum_{t=1}^{N} R_S(t). \tag{7}$$

To deal with the holiday effect we propose the concept of **one-day** returns $R_D(t)$. It is defined as the relative increase in price since the previous day, if that day is represented in the time series, i.e.

$$R_D(t) = 100 \cdot \delta(t) \cdot \frac{y(t) - y(t-1)}{y(t-1)}, \tag{8}$$

where

$$\delta(t) = \left\{ \begin{array}{l} 1 \text{ if } juliandate(t) - juliandate(t-1) = 1 \\ 0 \text{ otherwise} \end{array} \right\}. \tag{9}$$

The function *juliandate(t)* returns the number of days since the start of the Julian calender up to the date for sample t. It is introduced as a means to compute the number of days between two samples.

The mean of the **one-day** returns $R_D(t)$ is

$$R_D = \frac{1}{K}\sum_{t=1}^{N} R_D(t), \tag{10}$$

where

$$K = \sum_{t=1}^{N} \delta(t). \tag{11}$$

K is roughly the number of weekdays Tuesday-Friday in the data set. Holidays occurring on weekdays as well as missing data will also generate $\delta = 0$, and therefore a reduced value on K.

The standard deviations for one-day returns and one-step returns are computed correspondingly.

We first define increase and decrease fractions based on the standard one-step returns. The **one-step** Increase Fraction I_S is defined as

$$I_S = \frac{1}{K_S}\sum_{t=1}^{N} a_I(t), \tag{12}$$

where

$$a_I(t) = \left\{ \begin{array}{l} 1 \text{ if } y(t) > y(t-1) \\ 0 \text{ otherwise} \end{array} \right\} \tag{13}$$

and

$$K_S = \sum_{t=1}^{N} a_I(t) + a_D(t). \tag{14}$$

The **one-step** Decrease fraction D_S is defined as

$$D_S = \frac{1}{K_S} \sum_{t=1}^{N} a_D(t), \tag{15}$$

where

$$a_D(t) = \left\{ \begin{array}{l} 1 \text{ if } y(t) < y(t-1) \\ 0 \text{ otherwise} \end{array} \right\} \tag{16}$$

To deal with the holiday effect, we in similar fashion define increase and decrease fractions based on the one-day returns. The **one-day** Increase Fraction I_D is defined as

$$I_D = \frac{1}{K_D} \sum_{t=1}^{N} \delta(t) \cdot a_I(t), \tag{17}$$

where

$$K_D = \sum_{t=1}^{N} \delta(t) \cdot (a_I(t) + a_D(t)). \tag{18}$$

The **one-day** Decrease fraction D_D is defined as

$$D_D = \frac{1}{K_D} \sum_{t=1}^{N} \delta(t) \cdot a_D(t). \tag{19}$$

We start by presenting some simple statistics for the one-day and one-step returns. The results in Table 16 are mean values for the *SXG* stocks and the results in Table 17 are mean values for the *SXBIG* stocks. Results are presented for the years 1987-1996.

The results show a significant difference in both mean returns and variances for the two types of returns computed. The difference in the mean values are caused by a weekend effect and are investigated in more detail in the next section. The difference in variances show that the one-day returns exhibit a more homogenous distribution than the one-step returns. It can interpreted as a reduction in noise in the computation of $R_D(t)$ by removing the weekends, holidays and other missing data.

[3] The ratio between the two No.-of- obs columns should be 4/5 if only Friday-to-Monday returns had been removed. However, there are additional gaps in the data series originating from holidays and simply missing data.

Table 16: Average daily returns for 32 Swedish stocks (*SXG*)

One-Day-Returns $R_D(t)$			One-Step-Returns $R_S(t)$		
Mean	Std.dev.	No. of obs. [3]	Mean	Std.dev.	No. of obs. [3]
0.103	2.20	4558	0.077	2.29	5902

Table 17: Average daily returns for 207 Swedish stocks (*SXBIG*)

One-Day-Returns $R_D(t)$			One-Step-Returns $R_S(t)$		
Mean	Std.dev.	No. of obs. [3]	Mean	Std.dev.	No. of obs. [3]
0.132	3.25	18561	0.092	3.78	25924

Rejection of the trading-time hypothesis should call for some sort of special treatment when creating models for the prediction of future returns. However, this is seldom seen in published methods. Some researchers have noticed the problem, which seems to be even more prominent in the case of prediction of exchange rates. Weigend [7] deals with the situation by simply removing all but the Monday-to-Tuesday returns and performs the prediction task on this aggregated time series. Another way to deal with the problem would be to include the day of the week explicitly in the model.

5.2 Day-of-the-Week effect

The day-of-the-week effect has been studied in a number of research papers. Hawawini and Keim [3] present a summary, which demonstrates significant differences in average daily returns across days of the week. In our investigation of the Swedish stock market the daily returns are presented in a somewhat different fashion than is normally done. The stock returns are computed for all twenty-five combinations of buy and sell day. The returns are presented as "daily returns", i.e. they are divided by the number of calendar days between buy and sell. For example, the return from buying on Friday and selling on Monday is divided by three before it is put in the table. A second table with the same layout presents the "Increase Fraction" for the same combinations of buy and sell day. This number is the fraction of sampled buy/sell situations that resulted in a positive return (zero returns are removed before calculating the Increase Fraction). In this way all combinations of buy and sell day can be compared on an equal basis.

As before results are presented for the *SXG* set of stocks. A time period of 10 years (1987-1996) were used to generate the results. This provides statistically more stable grounds than using one single index as in the investigations presented in [3]. The daily returns R_D are shown in Table 18 and the Increase Fraction I_D in Table 19.

Table 18: Daily returns (%) for combinations of Buy and Sell days for SXG

	Sell Day					
Buy Day	Mon	Tue	Wed	Thu	Fri	**Mean**
Mon	0.045	0.005	0.033	0.078	0.088	0.050
Tue	0.062	0.056	0.094	0.138	0.133	0.097
Wed	0.054	0.048	0.057	0.174	0.159	0.099
Thu	0.023	0.018	0.033	0.049	0.138	0.052
Fri	-0.013	-0.012	0.002	0.027	0.046	0.010
Mean	0.034	0.023	0.044	0.093	0.113	0.061

Table 19: Increase fraction (%) for combinations of Buy and Sell days for SXG

	Sell Day					
Buy Day	Mon	Tue	Wed	Thu	Fri	**Mean**
Mon	50.58	48.99	49.88	51.20	51.67	50.46
Tue	51.13	52.05	50.73	52.42	52.61	51.79
Wed	51.38	51.75	52.83	53.08	52.86	52.38
Thu	49.20	49.72	50.87	51.61	51.27	50.53
Fri	48.41	48.02	48.99	50.24	51.43	49.42
Mean	50.14	50.11	50.66	51.71	51.97	50.92

The same type of tables as the above have also been generated for the *SXBIG* stocks. The daily returns R_D are shown in Table 20 and the Increase Fraction I_D in Table 21.

Table 20: Daily returns (%) for combinations of Buy and Sell days for SXBIG

	Sell Day					
Buy Day	Mon	Tue	Wed	Thu	Fri	**Mean**
Mon	0.069	0.004	0.050	0.078	0.105	0.061
Tue	0.081	0.066	0.114	0.136	0.152	0.110
Wed	0.076	0.060	0.066	0.167	0.186	0.111
Thu	0.061	0.044	0.051	0.058	0.243	0.091
Fri	0.014	0.003	0.018	0.033	0.062	0.026
Mean	0.060	0.035	0.060	0.095	0.150	0.080

Table 21: Increase fraction (%) for combinations of Buy and Sell days for SXBIG

	Sell Day					
Buy Day	Mon	Tue	Wed	Thu	Fri	**Mean**
Mon	50.21	48.42	49.07	49.89	50.66	49.65
Tue	50.85	50.68	50.39	51.24	52.04	51.04
Wed	50.79	50.27	51.16	51.63	52.40	51.25
Thu	49.80	49.38	49.86	50.41	52.05	50.30
Fri	48.70	47.51	48.18	49.02	50.30	48.74
Mean	50.07	49.25	49.73	50.44	51.49	50.20

We can extract several interesting "anomalies" from the last two tables:

- The day-of-the-week affects the returns in a significant manner. The returns span between 0.003% (buy Friday/sell Tuesday) and 0.243% (buy Thursday/sell Friday).
- The one-day returns increase monotonically from Monday to Thursday: 0.004, 0.114, 0.167, 0.243 (buying on a Friday never yields a one-day return).
- The rightmost column describes the mean returns achieved when selling between one and seven days from the buying day. Friday appears to be the worst day to buy in this respect.
- Looking at "Increase Fractions" it is still clear that the real trading conditions are almost as bad as before. Even if we pick the best choice and buy on Thursday and sell on Friday, we will loose money in 47.95% of the cases. It would take great patience and a stable financial backup to utilize the shown day-of-week effect.

A question that always should be posed when looking for and finding structures in huge data sets is whether the found structure reflects some general property of the data generating process, or it is simply an effect of data snooping. The general question is, at the least, hard to answer and is further discussed in [4]. In this particular case, we have calculated the same statistics for yearly data 1987-1996. In this way the results are tested for stability in time. The reported effects are present even in these cases and thus provide additional support for the results.

5.3 Month effects

The month effect on stock returns is investigated in a similar fashion as the day-of-the-week effect above. We perform two statistical investigations:

- Average daily returns for each month
- Average monthly returns for combinations of buy and sell month

Daily returns for each month. In this study we compute average daily returns $R_D(t)$ and $R_S(t)$ as defined in Section 5.1. The reason to use the "normal" $R_S(t)$ is the day-of-week effect shown in the previous two sections. By using the weekend compensated return $R_D(t)$ we achieve more stable figures with the "noise" caused by weekends, holidays and missing data removed. The returns are computed for the *SXG* and *SXBIG* stocks for the years 1987-1996. The mean results for *SXG* are shown in Table 22 and 23. The results for *SXBIG* are shown in Tables 24 and 25.

The results for the Swedish stock market fits well with investigations on other markets. Hawawini and Keim [3] present a summary of research on a number of stock markets world wide. The high returns for January and low returns for September are significant for most of the markets including the Swedish stock market.

Table 22: One-day returns for SXG

	Mean	**Std.dev.**	**Incr.fraction**	**No. of obs.**
Jan	0.355	2.37	55.92	4338
Feb	0.173	2.24	52.19	4310
Mar	0.065	1.95	48.62	4748
Apr	0.227	1.85	54.78	4130
May	0.133	2.06	52.22	4078
Jun	0.007	1.75	47.71	4426
Jul	0.194	1.61	55.86	4887
Aug	-0.002	2.17	48.17	5059
Sep	-0.028	2.37	49.72	4750
Oct	-0.019	2.73	48.33	5015
Nov	0.106	2.70	49.58	4872
Dec	0.076	2.18	50.52	4083
Mean	0.103	2.20	51.03	4558

Table 23: One-step returns for SXG

	Mean	**Std.dev.**	**Incr.fraction**	**No. of obs.**
Jan	0.331	2.48	55.07	5700
Feb	0.124	2.29	51.49	5499
Mar	0.034	1.98	48.31	6133
Apr	0.200	1.94	53.68	5486
May	0.178	2.08	53.17	5587
Jun	0.007	1.85	48.11	5709
Jul	0.176	1.71	54.68	6234
Aug	-0.105	2.27	46.29	6352
Sep	-0.038	2.49	48.65	6072
Oct	-0.102	2.88	48.22	6303
Nov	0.092	2.79	49.63	6170
Dec	0.081	2.32	50.99	5575
Mean	0.077	2.29	50.60	5902

Table 24: One-day returns for SXBIG

	Mean	**Std.dev.**	**Incr.fraction**	**No. of obs.**
Jan	0.376	3.32	54.44	17096
Feb	0.231	3.54	51.44	17307
Mar	0.019	2.93	47.80	19154
Apr	0.241	2.82	53.87	16224
May	0.143	3.02	51.28	16501
Jun	0.034	2.99	48.80	17750
Jul	0.237	2.64	54.10	18417
Aug	0.019	3.10	48.92	20524
Sep	0.041	3.38	50.03	19877
Oct	0.078	3.72	48.78	21181
Nov	0.133	3.52	49.58	20397
Dec	0.095	3.64	49.97	18303
Mean	0.132	3.25	50.62	18561

Table 25: One-step returns for SXBIG

	Mean	**Std.dev.**	**Incr.fraction**	**No. of obs.**
Jan	0.357	3.90	53.20	24058
Feb	0.208	4.03	50.76	23590
Mar	-0.025	3.43	47.35	26661
Apr	0.202	3.28	52.67	23408
May	0.152	3.47	51.62	24406
Jun	-0.026	3.23	48.10	24914
Jul	0.220	3.09	52.99	26042
Aug	-0.096	3.45	47.24	27941
Sep	-0.042	3.78	48.49	27338
Oct	0.036	4.57	48.77	28515
Nov	0.082	4.04	48.91	27764
Dec	0.111	4.56	50.33	26447
Mean	0.092	3.78	49.94	25924

Increase Fractions. The Increase Fraction I_D for *SXBIG* varies between 47.80% (March) and 54.44% (January). The mean I_D is 50.62%, which is close to the 50% proposed by the random walk hypothesis. Note that a prediction accuracy ("hit rates") of about 54% is often reported for elaborate prediction algorithms. Most algorithms do not use the day-of-week or the month-of-year as input variables, see e.g. [6] or [1]. These results are claimed to show predictive capability in the algorithms. Be that as it may, if we can achieve the same hit rate by just looking at what month we are trading in, it seems reasonable to incorporate in some way the month-of-year in the algorithm. And for algorithms that don't do that, the validation process should really be looked over.

Monthly returns for combinations of buy and sell month. We conclude the investigations of seasonal effects with a trading oriented statistical test, where both buying and selling are considered. The first trading day in each month is always selected for both buying and selling. After buying in the beginning of a month, the returns from selling in the beginning of each one of the successive twelve months are stored in a twelve-by-twelve table. The shown figures in this table are daily returns 30, to give comparable monthly returns for all months regardless of how many days they contain. A second table with the same layout shows the Increase Fraction for the same combinations of buy and sell month. This number is the fraction of sampled buy/sell situations that resulted in a positive return (zero returns are removed before calculating the Increase Fraction). In this way all combinations of buy and sell month can be compared on an equal basis.

Tables 26 and 27 show the results for the *SXBIG* stocks for the years 1987-1996. Tables 28 and 29 show the results for the *SXG* stocks for the same time period. To avoid false conclusions caused by extreme trends during parts of single years, the data is also investigated in finer detail. The figures 8, 9, 10, 11, 12 and 13 present results for the *SXG* stocks for the years 1987-1996. Each graph contains data for one specific buy month. Each point in the graphs shows the mean profit from selling all the stocks at that particular month in one particular year. Each point thus represents 32 trades. There are 10 points plotted for each sell month, one for each year between 1987 and 1996. Both buying and selling are done once per month on the first trading day of the month. The same analysis is done for the *SXBIG* set. The figures 14, 15, 16, 17, 18, and 19 present these results. The same combinations between buy and sell month are shown in the 3-D diagrams in Figure 7.

Table 26: Monthly returns (%) for combinations of buy and sell months for SXBIG

	Sell Month												
Buy Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	**Mean**
Jan	1.47	6.36	4.79	2.70	2.91	2.92	2.35	2.70	2.18	1.91	1.87	1.76	2.83
Feb	1.14	1.91	3.91	1.34	1.95	2.27	1.76	2.40	1.82	1.65	1.61	1.49	1.94
Mar	0.52	1.27	1.33	-0.58	1.12	1.60	1.10	1.77	1.22	1.01	1.00	0.94	1.03
Apr	0.73	1.52	1.51	1.12	3.31	2.80	1.70	2.46	1.59	1.27	1.23	1.13	1.70
May	0.35	1.25	1.27	0.87	1.18	2.56	0.90	2.07	1.09	0.78	0.82	0.73	1.16
Jun	-0.06	0.89	0.99	0.64	0.97	1.08	-0.65	1.49	0.51	0.27	0.26	0.29	0.56
Jul	0.16	1.20	1.32	0.85	1.19	1.28	1.13	4.01	1.11	0.62	0.57	0.54	1.16
Aug	-0.82	0.43	0.63	0.28	0.66	0.85	0.72	1.01	-1.53	-1.07	-0.68	-0.32	0.01
Sep	-0.49	1.06	1.24	0.76	1.13	1.31	1.15	1.54	1.28	-1.04	-0.45	0.08	0.63
Oct	-0.05	1.82	2.00	1.29	1.61	1.81	1.57	2.06	1.67	1.45	0.15	0.74	1.34
Nov	0.07	2.32	2.43	1.58	1.86	2.10	1.80	2.26	1.85	1.60	1.82	1.25	1.75
Dec	0.35	3.68	3.38	2.10	2.34	2.51	2.09	2.48	1.97	1.70	1.90	1.85	2.20
Mean	0.28	1.98	2.07	1.08	1.69	1.92	1.30	2.19	1.23	0.85	0.84	0.87	1.36

Table 27: Increase fraction (%) for combinations of buy and sell months for SXBIG

	Sell Month												
Buy Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	**Mean**
Jan	53.89	65.12	66.08	59.48	66.08	66.28	64.06	66.46	61.84	62.26	59.27	59.68	62.54
Feb	50.37	52.39	55.03	53.02	58.68	59.53	59.04	61.10	57.50	57.52	55.35	54.84	56.20
Mar	46.57	48.89	50.81	43.03	54.95	58.05	56.05	58.75	54.41	55.24	51.90	50.84	52.46
Apr	49.61	51.78	53.77	52.19	63.64	67.02	60.71	65.54	59.84	58.24	54.95	53.48	57.56
May	45.23	48.54	50.43	48.10	52.02	56.57	53.15	58.53	55.26	54.60	52.05	49.93	52.03
Jun	44.73	48.84	49.83	47.32	51.74	54.28	45.88	55.58	50.76	53.60	49.64	48.16	50.03
Jul	47.06	51.54	52.45	49.58	54.59	57.49	56.92	60.93	52.15	54.39	51.18	50.92	53.27
Aug	41.08	47.54	47.13	44.33	48.56	52.13	52.52	53.74	41.95	46.21	45.68	45.70	47.21
Sep	44.80	52.16	53.26	50.37	53.95	57.26	56.74	59.05	54.60	49.64	47.31	48.80	52.33
Oct	46.50	56.02	56.83	53.48	56.41	58.62	58.81	58.36	55.75	55.22	47.56	50.53	54.51
Nov	44.69	61.16	60.67	55.06	61.18	61.67	60.56	61.15	56.70	57.31	55.61	48.80	57.05
Dec	49.87	65.49	66.11	59.71	64.90	66.00	64.49	66.26	60.43	61.12	59.27	61.07	62.06
Mean	47.03	54.12	55.20	51.31	57.23	59.58	57.41	60.46	55.10	55.45	52.48	51.90	54.77

Table 28: Monthly returns (%) for combinations of buy and sell months for SXG

	Sell Month												
Buy Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	**Mean**
Jan	1.97	7.10	4.80	3.36	3.55	3.72	3.20	3.43	2.82	2.56	2.21	2.09	3.40
Feb	1.34	2.09	2.70	1.80	2.46	2.77	2.32	2.70	2.18	2.04	1.62	1.50	2.13
Mar	1.20	2.00	1.84	0.74	2.19	2.78	2.22	2.69	2.05	1.79	1.41	1.34	1.86
Apr	1.23	2.09	1.90	1.72	3.94	3.93	2.78	3.26	2.32	1.96	1.50	1.42	2.34
May	0.66	1.57	1.45	1.34	1.57	3.59	1.94	2.64	1.53	1.22	0.74	0.76	1.58
Jun	0.14	1.12	1.07	0.97	1.28	1.52	0.08	1.99	0.74	0.59	0.12	0.24	0.82
Jul	0.17	1.27	1.21	1.08	1.42	1.66	1.58	3.82	0.94	0.59	0.03	0.25	1.17
Aug	-0.85	0.51	0.59	0.52	0.92	1.23	1.16	1.48	-2.08	-1.17	-1.41	-0.68	0.02
Sep	-0.37	1.22	1.28	1.15	1.54	1.85	1.74	2.10	1.82	-0.92	-1.42	-0.30	0.81
Oct	0.26	2.18	2.18	1.87	2.22	2.52	2.33	2.72	2.33	2.13	-1.84	0.26	1.59
Nov	1.21	3.51	3.24	2.66	2.93	3.24	2.94	3.24	2.77	2.51	2.54	2.20	2.75
Dec	0.92	4.60	3.63	2.80	3.00	3.30	2.89	3.08	2.50	2.19	2.15	2.15	2.77
Mean	0.66	2.44	2.16	1.67	2.25	2.68	2.10	2.76	1.66	1.29	0.64	0.94	1.77

Table 29: Increase fraction (%) for combinations of buy and sell months for SXG

	Sell Month												
Buy Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	**Mean**
Jan	58.73	68.86	72.95	67.62	75.18	75.80	75.00	76.16	72.50	71.89	65.71	64.89	70.44
Feb	55.24	60.66	58.03	61.07	64.39	73.84	73.02	73.74	67.26	64.06	61.87	59.14	64.36
Mar	52.19	57.83	58.96	48.35	64.87	72.70	68.90	70.77	64.41	61.13	56.34	53.68	60.85
Apr	52.36	55.38	60.24	60.39	67.14	75.27	67.96	73.50	66.90	63.07	56.79	53.68	62.72
May	47.41	52.00	55.34	53.94	57.65	62.54	61.27	65.85	60.49	59.44	52.98	50.17	56.59
Jun	44.36	51.79	56.25	53.49	57.75	64.20	53.62	62.81	55.75	57.29	50.00	46.53	54.49
Jul	45.88	51.98	56.64	52.94	59.30	66.28	65.50	66.78	54.86	57.79	49.31	49.14	56.37
Aug	41.63	49.60	50.59	47.45	51.94	59.69	60.31	60.94	41.18	45.49	42.76	44.95	49.71
Sep	43.36	50.99	55.64	55.47	60.47	64.48	62.99	65.10	59.85	51.22	41.40	46.71	54.81
Oct	48.44	62.85	63.04	57.53	63.42	65.89	65.12	65.23	61.48	59.77	44.91	51.92	59.13
Nov	51.18	70.97	69.41	66.15	74.42	76.54	76.36	75.58	68.08	67.05	64.23	52.26	67.69
Dec	57.60	77.29	78.21	70.31	76.92	78.68	80.69	79.07	71.92	70.38	64.73	69.35	72.93
Mean	49.87	59.18	61.27	57.89	64.45	69.66	67.56	69.63	62.06	60.71	54.25	53.54	60.84

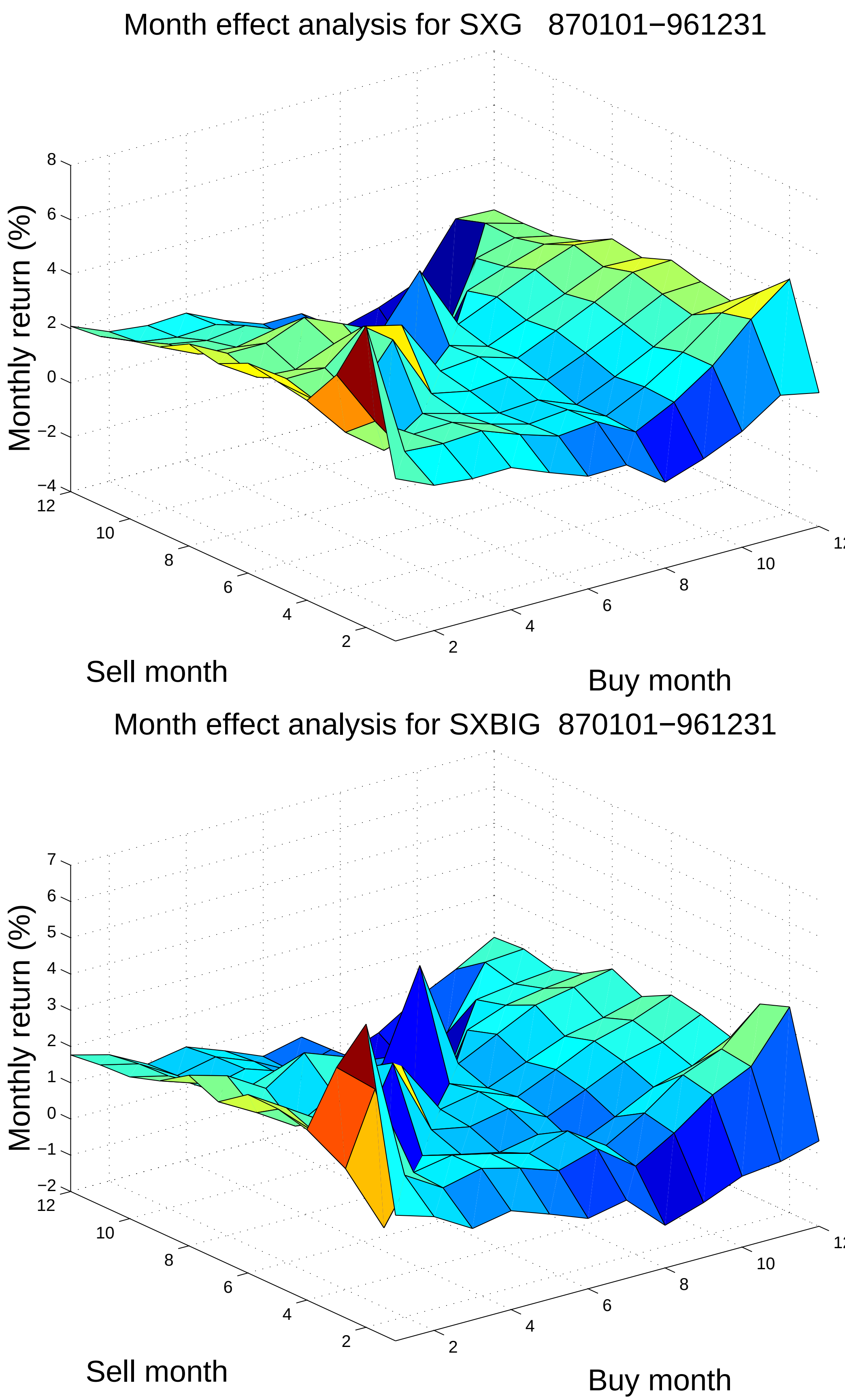

Figure 7: Monthly returns as a function of combinations of buy and sell month.

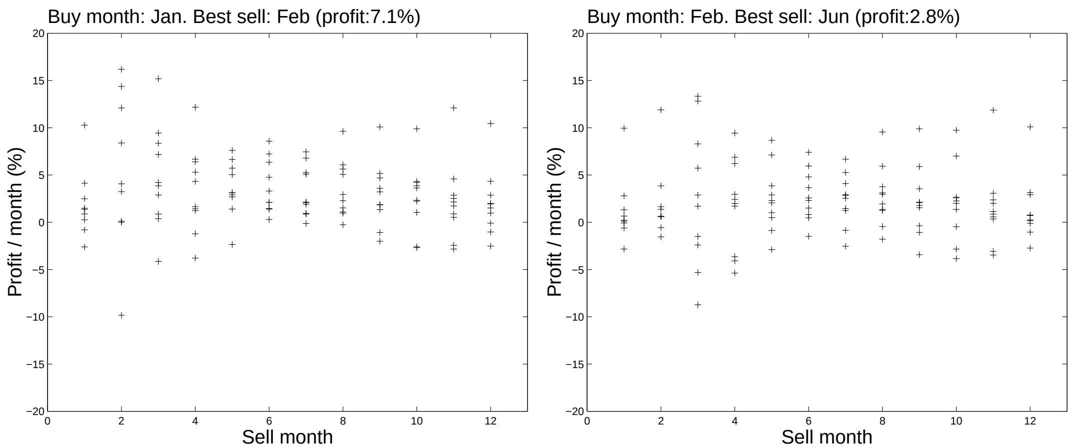

Figure 8: Monthly returns as a function of the sell month. Stocks: *SXG*. Years: 1987-1996.

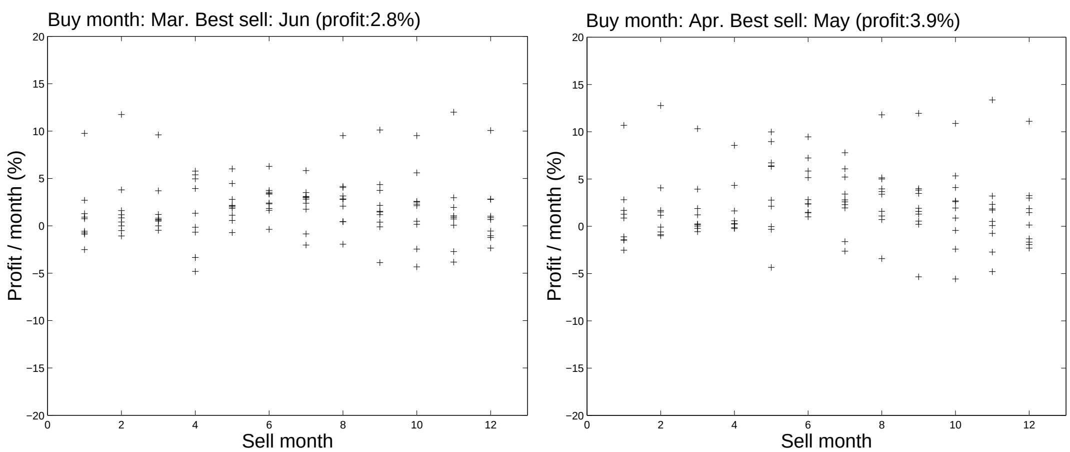

Figure 9: Monthly returns as a function of the sell month. Stocks: *SXG*. Years: 1987-1996.

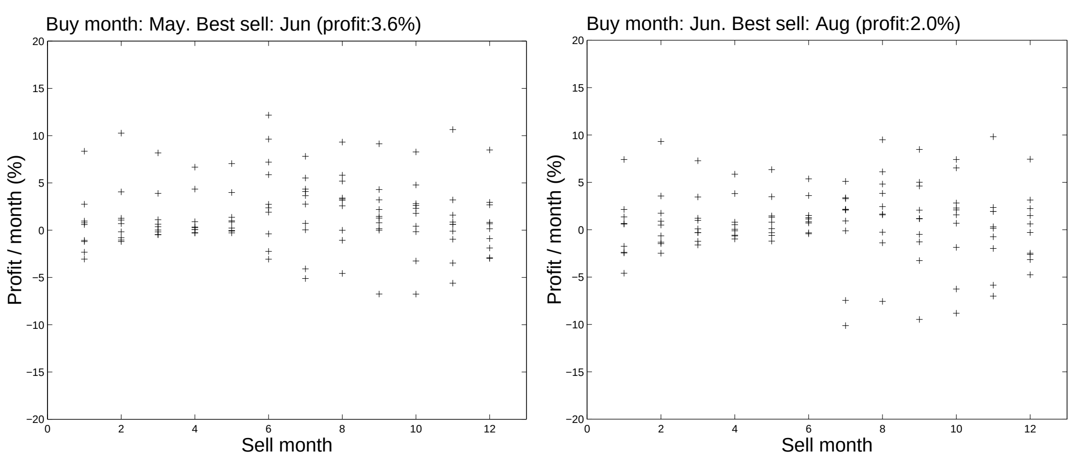

Figure 10: Monthly returns as a function of the sell month. Stocks: *SXG*. Years: 1987-1996.

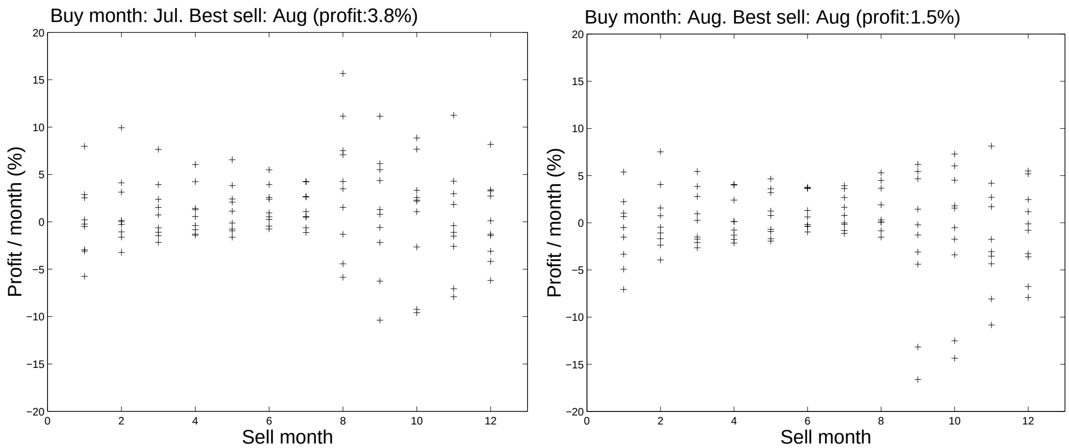

Figure 11: Monthly returns as a function of the sell month. Stocks: *SXG*. Years: 1987-1996.

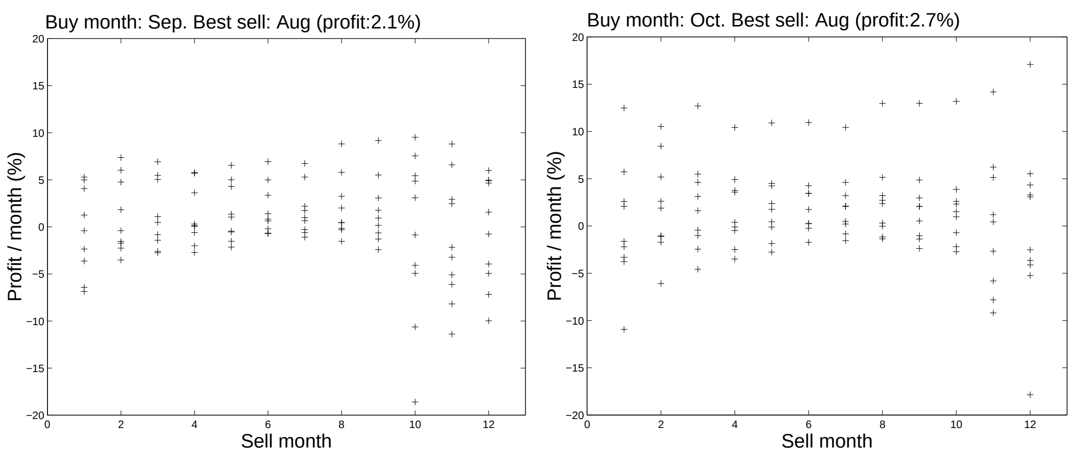

Figure 12: Monthly returns as a function of sell month. Stocks: *SXG*. Years: 1987-1996.

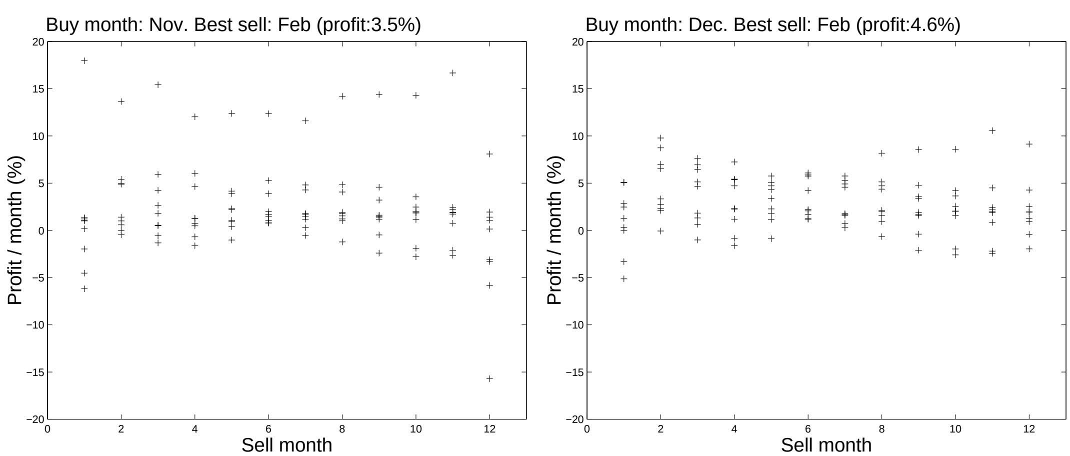

Figure 13: Monthly returns as a function of the sell month. Stocks: *SXG*. Years: 1987-1996.

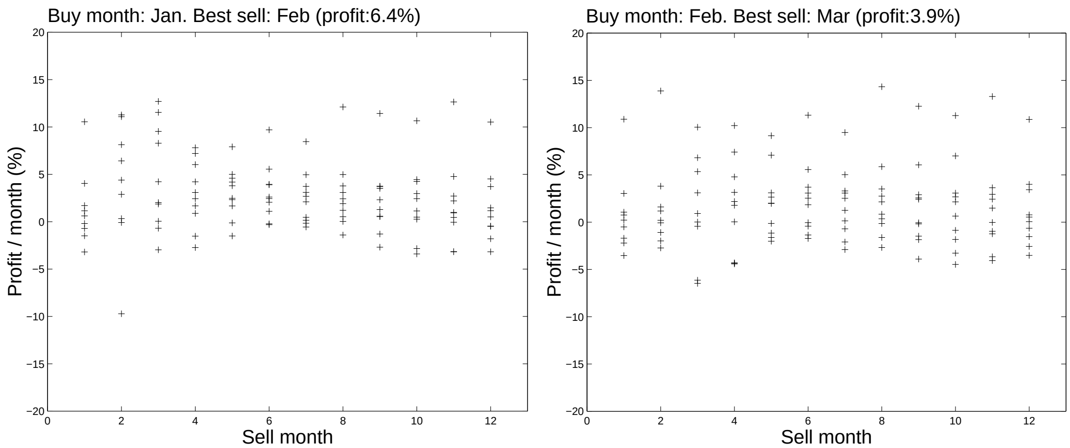

Figure 14: Monthly returns as a function of the sell month. Stocks: *SXBIG*. Years: 1987-1996.

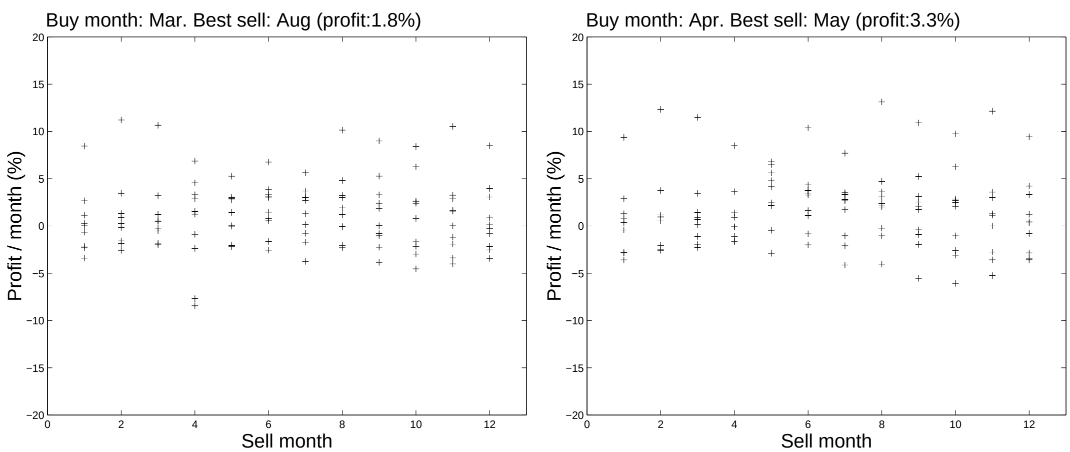

Figure 15: Monthly returns as a function of the sell month. Stocks: *SXBIG*. Years: 1987-1996.

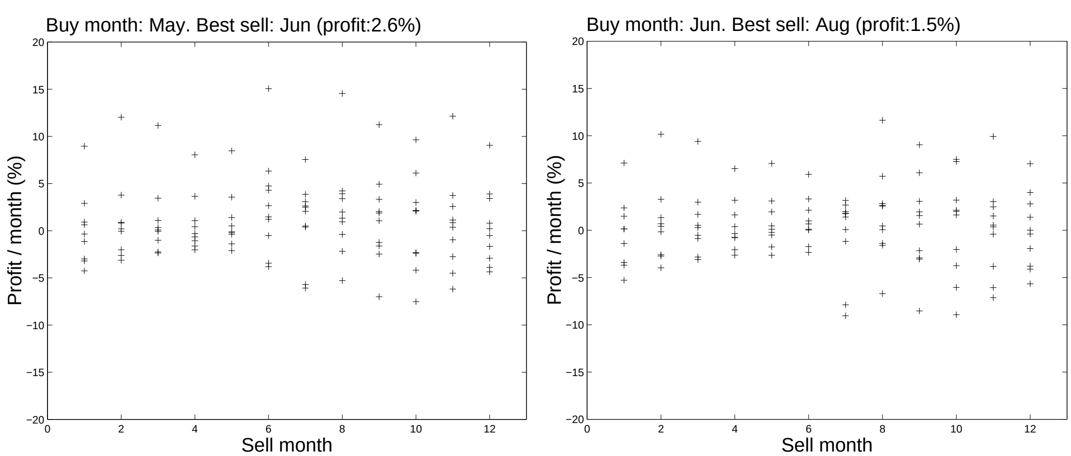

Figure 16: Monthly returns as a function of the sell month. Stocks: *SXBIG*. Years: 1987-1996.

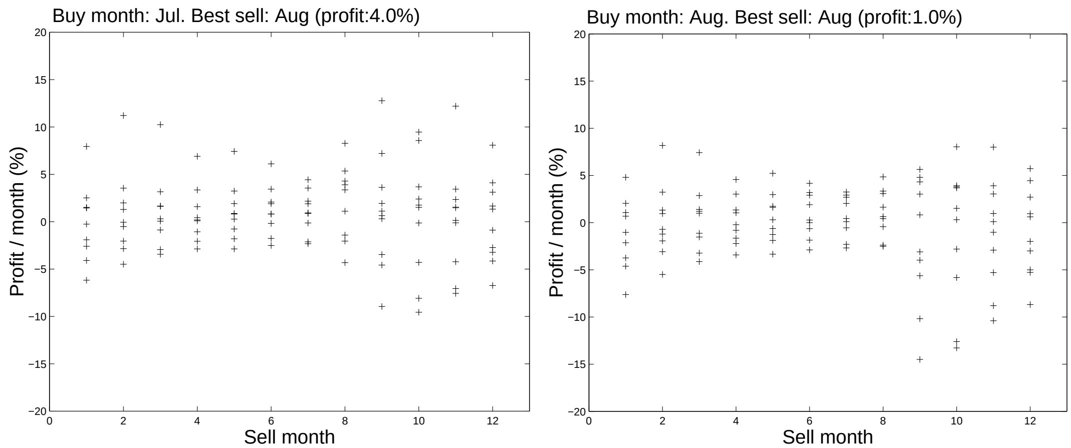

Figure 17: Monthly returns as a function of the sell month. Stocks: *SXBIG*. Years: 1987-1996.

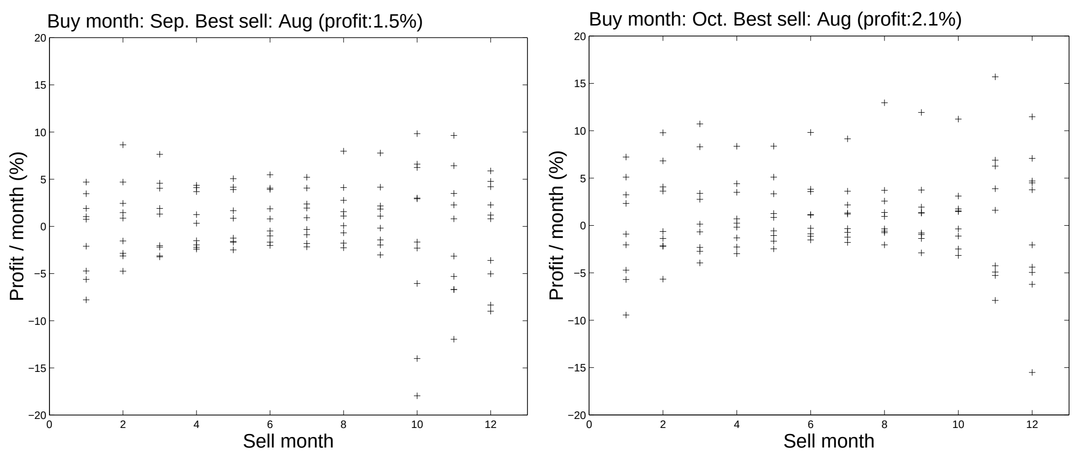

Figure 18: Monthly returns as a function of the sell month. Stocks: *SXBIG*. Years: 1987-1996.

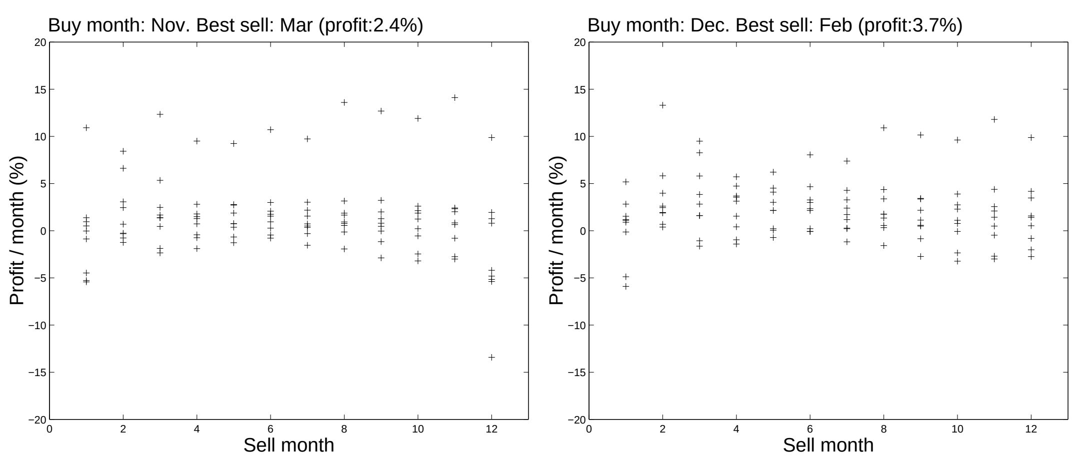

Figure 19: Monthly returns as a function of the sell month. Stocks: *SXBIG*. Years: 1987-1996.

6 Conclusions

- The investigations regarding autocorrelation show very low values with the possible exception of the first lag in the ACF. Surprisingly, the result for the 207 stocks in *SXBIG* shows a significantly *negative* correlation in the first lag, whereas *positive* correlation is found for the 32 major stocks in *SXG*.

- The analysis of trends give supporting results. We observe that a 5% decrease in price since yesterday, stands a 61.2% probability of showing an increase by tomorrow. The cases with large *increase* since yesterday exhibit a difference between the two investigated sets of stocks. For the stocks in *SXG*, a 5% increase in price since yesterday, stands a 54.9% probability of showing an *increase* by tomorrow. For the stocks in *SXBIG*, a 5% increase in price since yesterday, stands a 53.3% probability of showing a *decrease* by tomorrow.

  The massive better part of returns however falls into regions, where it is very difficult to claim *any* correlation between past and future price changes.

- The presented statistics show significant Day-of-the-week and Month effects on the stock returns. The proposed concept of daily returns R_D and a comparison to the ordinary step returns R_S showed a significant difference, caused by the holiday effect. The daily returns R_D vary between 0.004% (buy Monday/sell Tuesday) and 0.243% (buy Thursday/sell Friday). The Increase Fraction I_D was shown to depend on the month of the year and varies between 47.80% (March) and 54.44% (January). Even if the effects are too small to be utilized in actual trading, they are definitely big enough to influence other prediction algorithms such as ordinary time series analysis or neural network models of daily returns. If not taken into account in such algorithms, the seasonal effects will appear as a high noise level in data. It was shown that the Month effects are of the same size as the accuracy of many published prediction algorithms that don't make use of any date information.

  There are several ways to deal with the seasonal effects when constructing prediction algorithms:

  - Include the time dimension in the modeling, i.e. include a trainable parameter describing how the return depends on the day of the week or on the month.
  - Aggregate data. For example, instead of modeling the return time series for all days in a given time period, we can restrict the model to predict from one Monday to the next.
  - Without having done this statistical investigation, the most natural way to deal with the missing weekends might have been to expand the original time series $y(t)$ with interpolated data for Saturday and Sunday. However, the observed negative return between Friday and Monday shows that such an approach can not be recommended.

References

[1] D. J. E. Baestaens, W. M. van den Bergh, and H. Vaudrey. Market inefficiencies, technical trading and neural networks. In C. Dunis, editor, *Forecasting Financial Markets*, Financial Economics and Quantitative Analysis, pages 245–260. John Wiley & Sons, Chichester, England, 1996.

[2] E. F. Fama. Efficient capital markets: II. *The Journal of Finance*, 46(5):1575–1617, December 1991.

[3] G. Hawawini and D. B. Keim. On the predictability of common stock returns: Worldwide evidence. In R. A. Jarrow, V. Maksimovic, and W. T. Ziemba, editors, *Handbooks in Operations Research and Management Science*, volume 9: Finance, chapter 17, pages 497–544. North-Holland, Amsterdam, The Netherlands, 1995.

[4] T. Hellström and K. Holmström. Predicting the Stock Market. Research and Reports Opuscula ISRN HEV-BIB-OP–26-SE, Mälardalen University, Västerås, Sweden, 1998.

[5] D. Lowe and A. R. Webb. *Time series prediction by adaptive networks: A dynamical systems perspective.* IEEE Computer Society Press, 1991.

[6] G. Tsibouris and M. Zeidenberg. Testing the efficent markets hypothesis with gradient descent algorithms. In A.-P. Refenes, editor, *Neural Networks in the Capital Markets*, chapter 8, pages 127–136. John Wiley & Sons, Chichester, England, 1995.

[7] A. S. Weigend, B. A. Huberman, and D. E. Rumelhart. Predicting sunspots and exchange rates with connectionist networks. In M. Casdagli and S. Eubank, editors, *Nonlinear Modeling and Forecasting, SFI Studies in the Sciences of Complexity, Proc. Vol XII*, pages 395–432. Addison-Wesley, 1992.

[8] H. White. Economic prediction using neural networks: The case of IBM daily stock returns. In *IEEE International Conference on Neural Networks, San Diego*, pages 451–459, San Diego, 1988.