Can a simple model for the interaction between value and momentum traders explain how equity futures react to earnings announcements?

Johan Dellner

Stockholm, February 2011

Abstract

Hong and Stein (1999) explained the initial underreaction and the subsequent overreaction of prices to news as the outcome of the interaction between two groups of traders: news watchers and momentum traders. The news watchers have proprietary ways of interpreting public news and trade based on their interpretation. The true meaning of the news becomes gradually known to the crowd of news watchers and this creates the market underreaction. Underreaction makes momentum strategies profitable. Eventually, the momentum traders push the price too far and the market corrects. We test how well the model explains index and individual stock price behavior around earnings announcements. To remove ambiguity in the interpretation of the earnings news we proxy the news by the price change on the day of the announcement. Plots of the autocorrelation and the partial autocorrelation function suggest that the market reaction differs from that predicted by the model. There is an overreaction on the day of the announcement, a correction that lasts for 5-10 days and overshoots the price in the opposite direction and eventually a long trend with the same sign as the initial overreaction. To test the statistical significance of this observation we devise a trading strategy. Out-of-sample tests show some support for this observation. To explain the initial overreaction, presumably caused by very active momentum traders that trade during the announcement day, the model of Hong and Stein needs to include this group of traders and be applied on high frequency data during the announcement day.

Acknowledgements

I would like to thank Alex Gioulekas and Philip Nicolin at IPM for guidance during this thesis. I am also grateful for Professor Boualem Djehiche at KTH who was my tutor.

Stockholm, February 2011

Johan Dellner

Contents

1 Introduction

Hong and Stein (1999) proposed a model to explain underreaction, momentum trading and overreaction in asset markets. Their model assumes two groups of traders: news-watchers and momentum traders. Every news-watcher observes some private information, but fails to see the information given to other news-watchers. Thus news diffuses gradually across the population resulting in an underreaction of the asset price in the short run. This underreaction in the short run means that the momentum traders can profit from trend chasing strategies. However, if they only use a simple strategy, forecasting tomorrow's return based on yesterday's return, their actions will lead to an overreaction in the long run.

The existence of the news-watchers, who upgrade their views as they continuously receive more information, is motivated by evidence that stocks experience post-news drift in the direction the stock moved on the day of the news release, see Bernard (1992). Our findings suggest that this picture can be embellished by an initial short-term move in the opposite direction.

Commodity Trading Advisors (CTAs) are hedge funds that use momentum as their trading tool. These strategies are profitable as returns tend to exhibit positive correlation at three to twelve months' time horizons, see Jegadeesh and Titman (1993). DeBondt and Thaler (1985) find negative correlation between stock returns at time horizons beyond twelve months: at some point the trend loses steam and the price corrects towards its fundamental value.

The rest of the thesis is organized as follows. In Chapter 2, we explain how the model is constructed. We first introduce the news-watchers and show that the diffusion of information to these traders leads to underreaction. Next, we add momentum traders to the model. These traders pick up the price change and arbitrage away the underreaction left behind by the news-watchers. However, since the momentum traders use a simple trading strategy they create an overreaction.

In Chapter 3, we explain with stylized examples how to determine the model parameters. We find that the parameters in the model can be determined by different correlations plots.

In Chapter 4, we apply the model to reality. We produce the key correlation plots for stock indices and individual stocks. These plots do not resemble the plots of the stylized examples. This indicates that the model does not describe reality accurately. A different pattern emerges: the momentum traders are smart enough to not only condition on price change but also on the date of the news releases. They trade actively on the announcement day and cause the price to overreact. Hong and Stein (1999) mention that their model is intended to describe the price dynamics in response to private news. Then the momentum traders have no idea whether they are buying early in the cycle (generating profit) or late in the cycle (making a loss). When the news is public, the momentum traders are smart enough to refine their strategy: they make their strategy time-dependent and trade aggressively in the period just after the public announcement. To test the validity of this observation, we develop a trading strategy and test its performance out of sample. This trading strategy notes the sign of the return on the day of the news release, takes the opposite position for the next five days and reverses the position and keeps it for the following year.

2 The Model Construction

We start the model construction with the news-watchers, who trade a risky asset paying a single dividend at some later time T, the ultimate value of the dividend $D_T = D_0 + \sum_{j=0}^{T} \epsilon_j$, where ϵ_j is i.i.d normally distributed with mean zero and variance σ^2. Here we make the assumption that every ϵ_j can be decomposed into z independent parts, each with the same variance. At time t, information about ϵ_{t+z-1} begins to spread and has at this time been seen by a fraction $\frac{1}{z}$ of the total group of news-watchers. At the later time $t+1$ the information about ϵ_{t+z-1} has been seen by a fraction $\frac{2}{z}$ of the news-watchers, at time $t+2$ it has been seen by a fraction $\frac{3}{z}$ and so forth. This continues until ϵ_{t+z-1} has been seen by everyone, which happens at time $t+z-1$. The parameter z can be interpreted as the rate of information flow; a high value of z indicates a slow diffusion while a low value indicates a more rapid diffusion. Given this setup the price at time t, becomes;

$$P_t^{News-watchers} = D_t + \frac{z-1}{z}\epsilon_{t+1} + ... + \frac{1}{z}\epsilon_{t+z-1}, \qquad (1)$$

Next, we add the momentum traders to the model. We assume that at time t a momentum trader takes a position, which he holds for exactly j periods, until time $t+j$. Momentum traders submit quantity orders; the price is then determined by the competition against the news-watchers. They try to predict $(P_{t+j} - P_t)$ to determine the size of their orders. They use a simple univariate forecasting strategy only looking at the previous price change $\Delta P_{t-1} \equiv P_{t-1} - P_{t-2}$. One could allow the momentum traders to use n lags of price changes instead and give a different weight to each lag n. This would be a more realistic model of the behaviour of the trend-followers. However, we use the simplest model possible and assume that they do not have the computational horsepower to run a complicated multivariate strategy. So each momentum trader has an order flow of F_t at time t;

$$F_t = \phi \Delta P_{t-1},$$

where ϕ is an elasticity parameter and he holds this position until time t+j. The demand from momentum traders added together with the demand from news-watchers results in the following price at time t;

$$P_t = D_t + \frac{z-1}{z}\epsilon_{t+1} + ... + \frac{1}{z}\epsilon_{t+z-1} + \sum_{i=1}^{j} \phi \Delta P_{t-i}, \qquad (2)$$

Figure 1 shows that the momentum traders decrease the time of the underreaction in the beginning and trigger the overreaction.

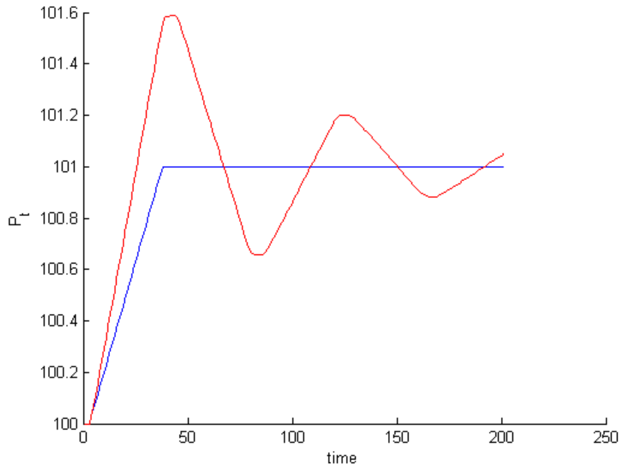

Figure 1: How the market reacts to one piece of news. The blue line represents the model with only the news-watcher (equation 1), while in the red line the momentum traders have been added (equation 2).

This model of P_t results in that ΔP_t can be written in the following way;

$$\Delta P_t = P_t - P_{t-1} =$$

$$D_t + \frac{z-1}{z}\epsilon_{t+1} + ... + \frac{1}{z}\epsilon_{t+z-1} + \sum_{i=1}^{j} \phi \Delta P_{t-i}$$

$$-(D_{t-1} + \frac{z-1}{z}\epsilon_t + ... + \frac{1}{z}\epsilon_{t+z-2} + \sum_{i=1}^{j} \phi \Delta P_{t-i-1}) =$$

$$= D_0 + \epsilon_0 + ... + \epsilon_t + \frac{z-1}{z}\epsilon_{t+1} + ... + \frac{1}{z}\epsilon_{t+z-1} + \phi P_{t-1} - \phi P_{t-j-1}$$

$$-(D_0 + \epsilon_0 + ... + \epsilon_{t-1} + \frac{z-1}{z}\epsilon_t + ... + \frac{1}{z}\epsilon_{t+z-2} + \phi P_{t-2} - \phi P_{t-j-2}) =$$

$$= \frac{\sum_{i=0}^{z-1} \epsilon_{t+i}}{z} + \phi \Delta P_{t-1} - \phi \Delta P_{t-(j+1)},$$

This is an ARMA(p,q) model, where p=j+1 and q=z-1.

3 Stylized examples

In order to increase our understanding of this model, we investigate some stylized examples. We do this in several stages. In the first stage we examine the reaction of only the news-watchers part when the news appears once. We then consider news that arrives every twenty days and finally when news appears daily. We find how to determine the parameter z. Finally, we add the momentum traders and investigate how to identify the parameters j and z.

3.1 News-watchers

With only the news-watchers, the return $\Delta P_t^{News-watchers}$ turns out to be a MA(z-1)-process since;

$$\Delta P_t^{News-watchers} = P_t^{News-watchers} - P_{t-1}^{News-watchers} =$$

$$= D_t + \frac{z-1}{z}\epsilon_{t+1} + ... + \frac{1}{z}\epsilon_{t+z-1} - (D_{t-1} + \frac{z-1}{z}\epsilon_t + ... + \frac{1}{z}\epsilon_{t+z-2}) =$$

$$= \frac{\sum_{i=0}^{z-1}\epsilon_{t+i}}{z},$$

A well-known technique (see e.g. Chapter 3 in Brockwell and Davis (1991)), to identify the parameter q in a MA(q)-process is to look at the autocorrelation function. The autocorrelation is defined as;

$$Autocorrelation\ at\ lag\ i \equiv Corr(\Delta P_k, \Delta P_{k+i}) \qquad k = 1, 2, 3, ..., L$$

where the Corr stands for the correlation which is defined as;

$$Corr(\Delta P_k, \Delta P_{k+i}) \equiv \frac{cov(\Delta P_k, \Delta P_{k+i})}{\sqrt{Var(\Delta P_k)}\sqrt{Var(\Delta P_{k+i})}}$$

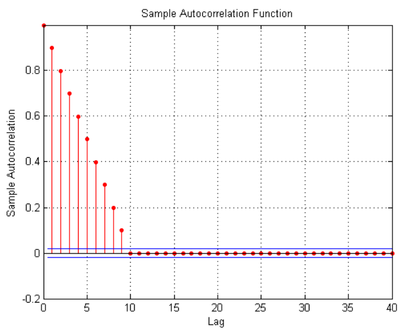

Figure 2: The autocorrelation function is plotted with only the news-watchers when the news appears once, all ϵ_i is equal to zero except one, with the parameter z set equal to 10.

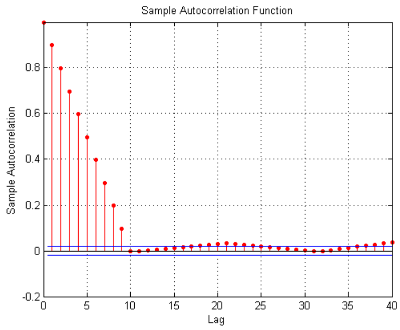

Figure 3: The autocorrelation function is plotted with only the news-watchers when news arrives every twenty days, every twenty ϵ_i is normal distributed $N(0,1)$, with the parameter z set equal to 10.

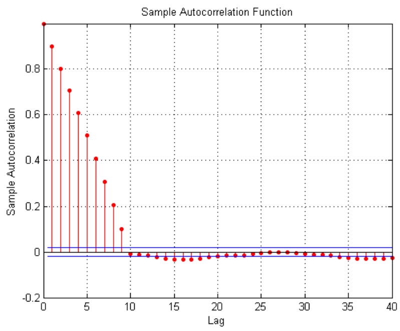

Figure 4: The autocorrelation function is plotted with only the news-watchers when news appears daily, every ϵ_i is normal distributed $N(0,1)$, with the parameter z set equal to 10.

In Figure 2, Figure 3 and Figure 4 the last positive value in the auto-correlation plots are at 9, with z equal to 10. This is in accordance with our model, since news is being spread until time $z-1$, thus it should be a positive and decreasing correlation up to that point.

3.2 News-watchers and Momentum traders

With the momentum traders added, the model is not a MA(z-1)-process anymore; instead it is an ARMA(j+1,z-1) model. From statistical theory (see e.g. Chapter 3 in Brockwell and Davis (1991)) we know that the parameter p in a AR(p)-process is determined from the partial autocorrelation function. The partial autocorrelation function is defined as;

$$\phi_{kk} = Corr(X_t - P(X_t|X_{t+1}, ..., X_{t+k-1}), X_{t+k} - P(X_{t+k}|X_{t+1}, ..., X_{t+k-1})),$$

where, $P(W|Z)$ is the best projection of W on Z. The interpretation of this is that it is the autocorrelation between X_t and X_{t+k} with the linear dependence from X_{t+1} to X_{t+k-1} removed.

In a causal auto regression model, AR(p);

$$X_t = Z_t + \phi_1 X_{t-1} + ... + \phi_p X_{t-p}, \quad Z_t \sim WN(0, \sigma^2),$$

the partial autocorrelation is zero for all lags k when k>p. By definition;

$$P(X_{k+1}|X_2, ..., X_k) = \sum_{j=1}^{p} \phi_j X_{k+1-j},$$

if Y is a linear combination of $\{X_2, ..., X_k\}$, then by causality Y is a linear combination of $\{Z_j, j \leq k\}$, and

$$\langle X_{k+1} - \sum_{j=1}^{p} \phi_j X_{k+1-j}, Y \rangle = \langle Z_{k+1}, Y \rangle = 0,$$

and this implies that

$$\phi_{kk} = Corr(X_{k+1} - \sum_{j=1}^{p} \phi_j X_{k+1-j}, X_1 - P(X_1|X_2, ..., X_k)) =$$

$$= Corr(Z_{k+1}, X_1 - P(X_1|X_2, ..., X_k)) = 0,$$

3.2.1 News appears once

We choose the parameters j, z and ϕ in an arbitrary way but make sure that the model remains stable (a large value of ϕ would result in unstable and oscillating time series). They are set to z=5, j=20 and ϕ=0.37 at this initial stage. The autocorrelation of the returns is plotted in Figure 5 and the partial autocorrelation in Figure 6. It is not that easy to see the size of z as it was for the MA(q)-process in the autocorrelation plot. Neither can the j parameter be determined by the partial autocorrelation plot, which was the case in a pure AR(p)-process. However, it turns out that the autocorrelation of returns contains all the necessary information. The autocorrelation plot has a decreasing positive correlation until z-1, like the pure MA-process. However, in addition to the pure MA-process, we have an auto regression part clearly seen for the first time at j+1-z and peaking at j+1.

How does the situation look if $j = 5$, $z = 20$ while ϕ remains the same? Now the autocorrelation function (see Figure 7) fails to give us the information it gave us in the previous case. This is due to the fact that the two processes, the AR(p) and MA(q) gets "mixed up" with each other. A way to understand this is to think what the first case with z=5 and j=20 actually means in our model. It means that the news are completely spread at time t=5 and that simplifies things since the effect of j=20 ($-\phi \Delta P_{t-(j+1)}$) first come into effect at t=j-z=15. While in the other case with $j = 5$, $z = 20$ both the momentum traders and the news-watchers are trading at the same lags, which complicates the picture of the autocorrelation function. It turns out that the partial autocorrelation function holds the information about j and z. However, it is not as simple as in the previous case.

From a rigorous inspection of the plots of the partial autocorrelation, it is clear that for our model the partial autocorrelation takes the value

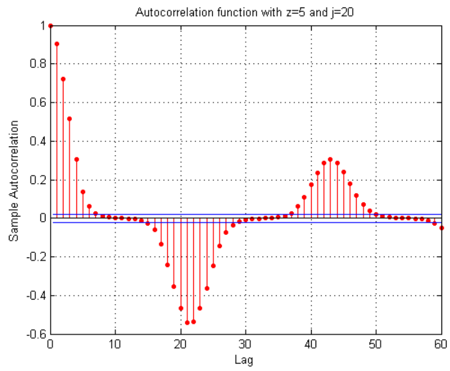

Figure 5: The autocorrelation of an impulse signal of one at time=1, when z=5, j=20 and ϕ=0.37

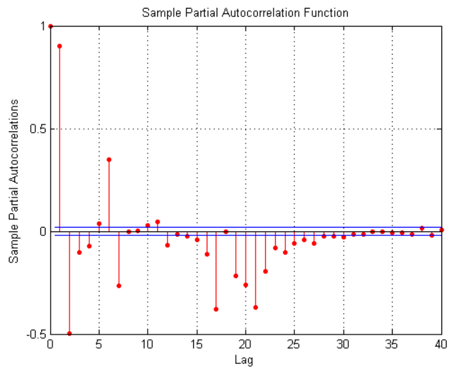

Figure 6: The partial autocorrelation of an impulse signal of one at time=1, when z=5, j=20 and ϕ=0.37.

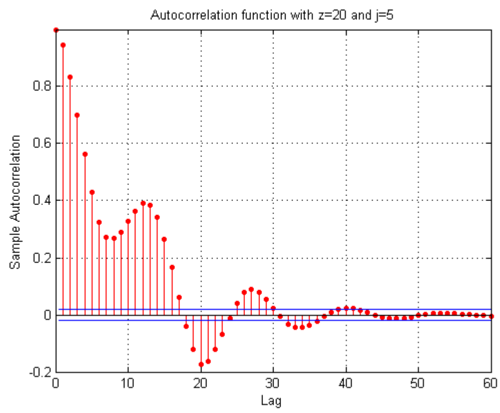

Figure 7: The autocorrelation of an impulse signal of one at time=1, when z=20, j=5 and ϕ=0.37

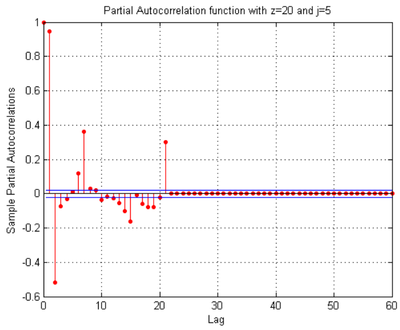

Figure 8: The partial autocorrelation of an impulse signal of one at time=1, when z=20, j=5 and ϕ=0.37

zero after z+1 lags and has a positive peak at j+2 if and only if $z \geq j$. We conclude that if $2z \leq j$ the parameters can be determined from the autocorrelation plots while if $z \geq j$ the parameters can be determined by the partial autocorrelation. However, the values when $2z > j > z$ remains uncertain. We will return to this case later.

3.2.2 News arrives every twenty days

When the news-watchers receive news every twenty days, every ϵ_{20k} for every integer k will be normally distributed with mean 0 and variance 1. This is illustrated in a plot of the price with and without the news-watchers in Figure 9.

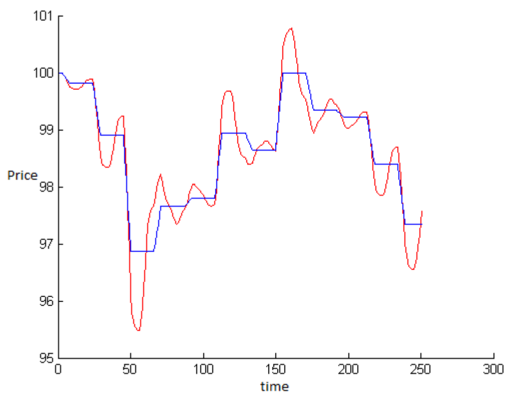

Figure 9: How the price responds from receiving news every 20 days, in this case $j = 20$, $z = 5$ and $\phi = 0.37$. The blue line is without the momentum traders and the red line is with them.

In the same way as with the impulse signal, we plot the autocorrelation function in Figure 10 in order to determine j and z. The plot is almost identical to the one with the impulse signal.

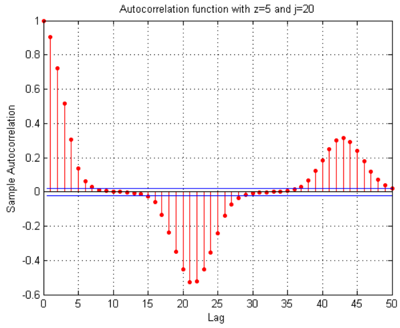

Figure 10: The autocorrelation function when news appears every twenty days, with $z = 5$, $j = 20$ and $\phi = 0.37$.

Then, we switch z and j in the same way as in the previous section so that j=5 and z=20. This is illustrated in Figure 11. The plot is almost identical to the one with the impulse signal. However, when it comes to the partial autocorrelation, there are some differences. Instead of turning zero after z+1, it begins to repeat itself with a decreasing factor for every period. Every period is exactly z lags long and there is a peak at j+2 in the same way as for the impulse signal.

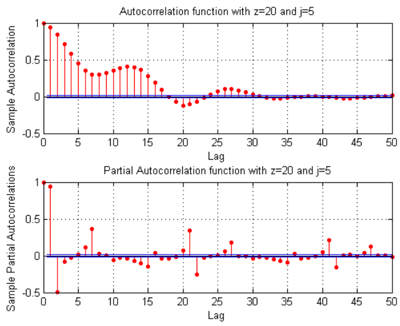

Figure 11: The autocorrelation and the partial autocorrelation function when news appears every twenty days, with z=20, j=5 and ϕ=0.37.

How do we determine z and j if $2z > j > z$? We assume that ϵ_{20i} are released at day $20i$ for all integers i. We plot the correlation

$$Correlation\ from\ news\ distributed\ at\ time\ 20i = Corr(\epsilon_{20i}, \Delta P_{20i+p}),$$

where p is the numbers of lags. Figure 12 shows this correlation when z=15 and j=20. The plot can intuitively be understood as follows. While the news is still being spread the correlation remains high. But once the news is completely spread, the correlation falls. However, it does not fall below zero.

The equation for ΔP_t shows the $\phi \Delta P_{t-1}$ will still give a positive result if no other term is involved. At time point $j+2$ it will fall to a negative value, since $-\phi \Delta P_{t-(j+1)}$ will react to what happened at time point 1. From this analysis the parameters z and j can be determined when $2z > j > z$.

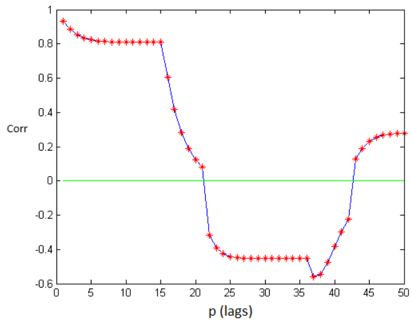

Figure 12: The correlation between the news (ϵ_{20i}) and the price change (ΔP_{20i+p}), when z=15, j=20 and ϕ=0.37.

3.2.3 News appears daily

We compute the same plots when news appears daily, every ϵ_t is normally distributed with mean 0 and variance 1. The analysis is the same as when news arrives every twenty days, see Figure 13-15

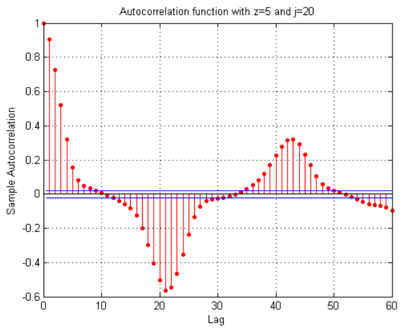

Figure 13: The autocorrelation function when news appears daily, with z=5, j=20 and ϕ=0.37.

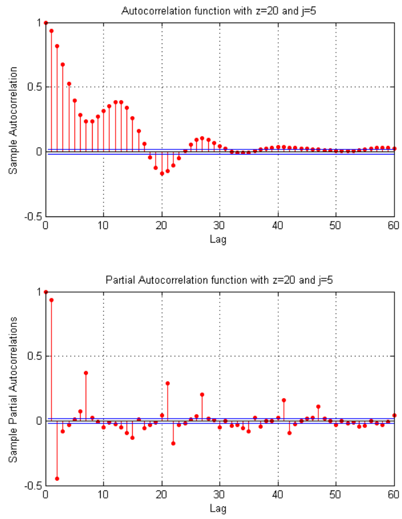

Figure 14: The autocorrelation and the partial autocorrelation function when news appears daily, with z=20, j=5 and ϕ=0.37.

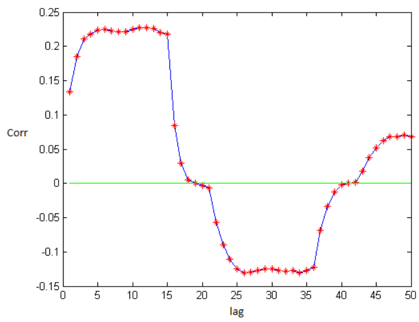

Figure 15: The correlation between the news (ϵ_i) and the price change (ΔP_{i+p}), when z=15, j=20 and ϕ=0.37.

3.3 Summary of the stylized example

When the news-watchers are the only group of traders in the model, the z parameter is easy to identify by looking at the autocorrelation plot. However, when momentum traders are added, it becomes harder to identify the parameters from the autocorrelation plot, especially for $2z > j$. When $z \geq j$ an additional plot is required, namely the partial autocorrelation plot. When $2z > j > z$ a correlation plot between the actual news and the price change the days following the news announcement is necessary. The methodology for identifying the model parameters is independent of the frequency of the news release.

4 Application to real market data

4.1 What constitutes news?

There are two main types of news: private news and public news. The public news, e.g. earnings announcements, is simultaneously observed by all investors. Private news is only observed by a fraction of the investors. The interpretation of the model is that the ϵ represents private news gradually diffusing across the population. When we apply the model to real market data, we are forced to use public news as our ϵ. Hong and Stein (1999) argue that even if the announcement itself is public, private information and judgment are required to evaluate the announcement.

When the news is public, smart momentum traders refine their strategy, they make their strategy time-dependent and trade aggressively in the period just after the public announcement. They exploit the profitable early stages of the trend. However, for now we will assume that the momentum traders are not that sophisticated. We will see that price patterns suggest that momentum traders are smart indeed.

How do we judge an earnings announcement? In previous papers, two different methods have been used to establish the market surprise of earnings announcement. Livnat and Mendelhall (2006) estimate the market surprise as the actual earnings minus the mean of the past six earnings figures. The idea behind this method is that past announcements in some sense are a good estimate of what will happen next. Chang, Jegadeesh and Likonishok (1996) look instead at how the market reacts on the day of the earnings announcement and take the price change that day as their market surprise. We follow this approach here, since we think that the best measurement of market surprise is the price reactions. The downside with using this method is that we cannot apply the model to the announcement day, we can apply it only to the following days. So our news can be written as;

$$\epsilon_{t+i} = s\delta_0(i)\Delta P_{t+i-z}, \qquad i = 0, 1, ..., z-1,$$

where $\delta_0(i)$ indicate a Dirac delta function which takes the value zero if no news has been published the period i otherwise one and s is a scaling factor.

4.2 Equity futures

Index return data for the last 15 years are retrieved from Bloomberg for four different markets: the S&P 500 (SPX, USA), the FTSE (UKX, United Kingdom), the Topix (TPX, Japan) and the ESTOXX (SXXE , Europe).

4.2.1 Autocorrelation and partial autocorrelation functions

If our model were true, the parameters z and j would be identifiable in the same way as in our stylized example. At first, we plot the autocorrelation function and the partial autocorrelation function for all the four markets, see Figure 16-19. These plots show none of features we observed in our stylized examples. This indicates that our model cannot capture the reality accurately. We will look at possible reasons for this.

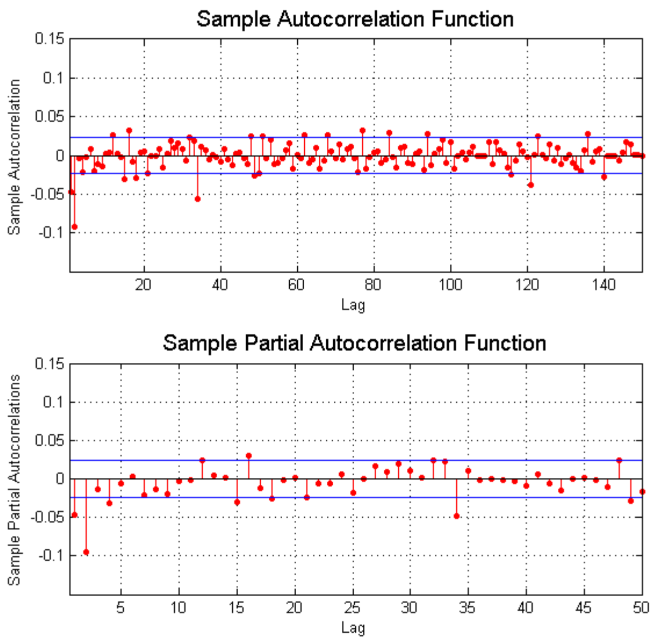

Figure 16: The autocorrelation and the partial autocorrelation function plots for the SPX daily returns.

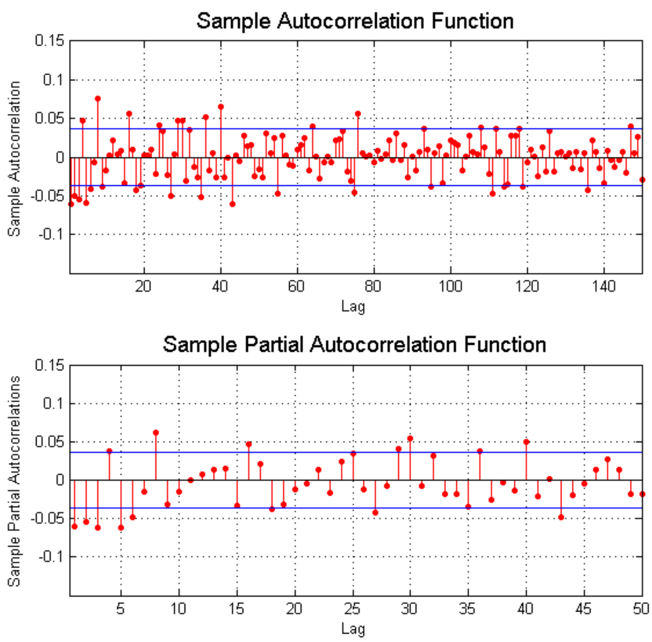

Figure 17: The autocorrelation and the partial autocorrelation function plots for the SXXE daily returns.

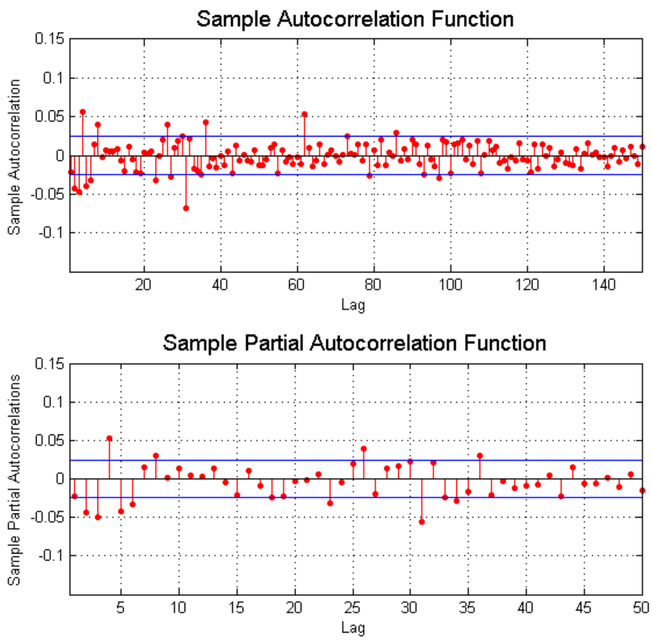

Figure 18: The autocorrelation and the partial autocorrelation function plots for the UKX daily returns.

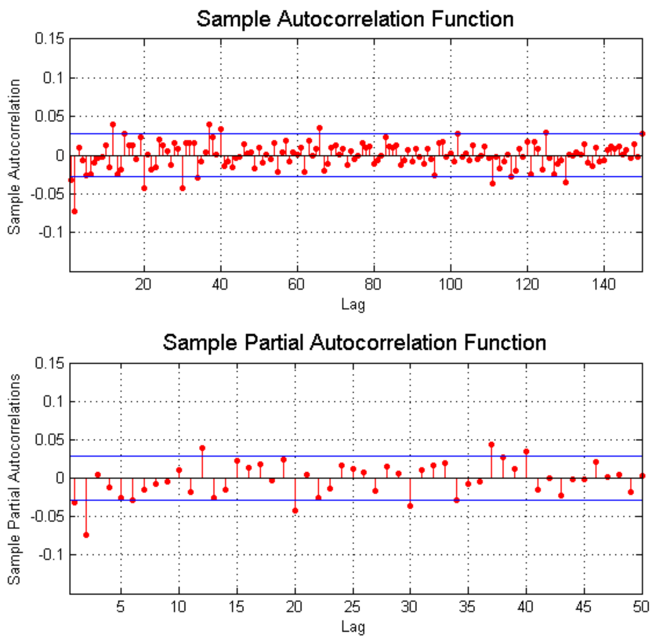

Figure 19: The autocorrelation and the partial autocorrelation function plots for the TPX daily returns.

4.2.2 Correlation with the news

We retrieved from Bloomberg the dates of the earnings announcements for each company in the index for the last quarter of 2010. We then took the earnings season to be the period with the highest concentration in earnings announcements. These periods span between 4-6 weeks and are presented in Table 1. As noted in the table, the companies of SPX are the first to announce, while announcements in the three remaining indices lag by a few weeks. Finally we assumed that quarterly earnings seasons have been recurring over the past 15 years. This means that the SPX earnings seasons were from the second Monday of January to the second Friday of February, from the second Monday of April to the second Friday of May, from the second Monday of July to the second Friday of August and from the second Monday of October to the second Friday of November every year. During these periods the Dirac delta function takes the value one, otherwise zero in:

$$\epsilon_{t+i} = \delta_0(i)\Delta P_{t+i-z}, \qquad i = 0, 1, ..., z-1,$$

Index	Start date	End date
SPX	Second Monday of January	Second Friday of February
SXXE	Fourth Monday of January	First Friday of March
UKX	Last Monday of January	Last Friday of February
TPX	Fourth Monday of January	Second Friday of February

Table 1: The period of when the earnings announcements are for the different markets for the last quarter of 2010.

We then plot the correlation between the news and the price change the following days, $Corr(\epsilon_t, \Delta P_{t+p})$ for every t, where p is the lag, see Figure 20-23. These plots are very different from the ones presented in the stylized example and look like plots of noise. However, if we fix t to be the announcement day and look at the correlation between the news (i.e. the return on the announcement day) and the return over the next p days,

$$Corr(\epsilon_t, P_{t+p} - P_t)$$

A certain pattern emerges: Figures 24 - 27 show that the correlation is negative during the first 5-10 days and then it increases.

This can be interpreted as follows. Five to ten days after the news announcement, the price moves in the opposite direction of the announcement return. After this initial period, the price enters a trend in the direction of the announcement return.

This is not in accordance with the model formulated in chapter 2. Instead of an underreaction to the news and a slow oscillation towards the "fundamental value", these plots indicate a price reaction similar to the illustration in Figure 28. We will test via a trading strategy whether this behaviour is statistically significant.

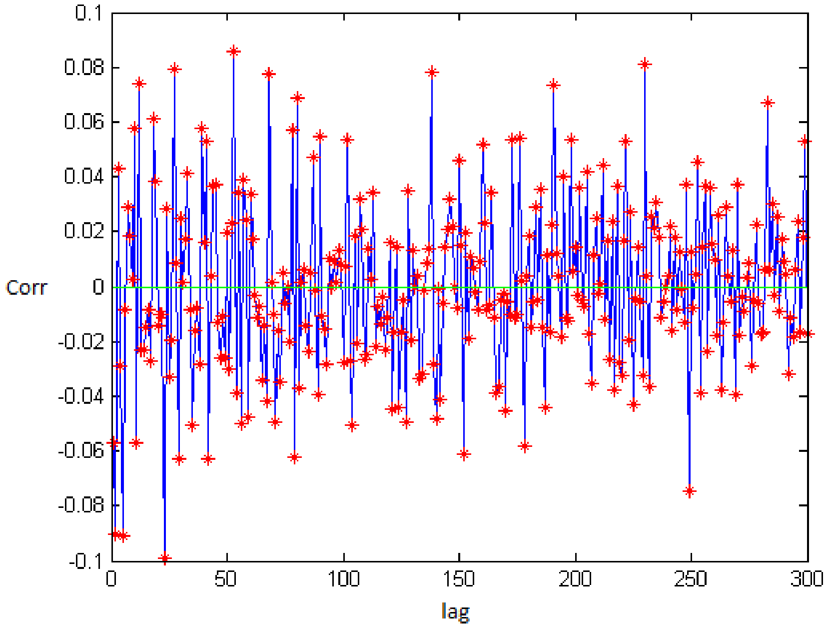

Figure 20: The correlation with the news ϵ and the price change lag p days after ΔP_{t+p} for SPX.

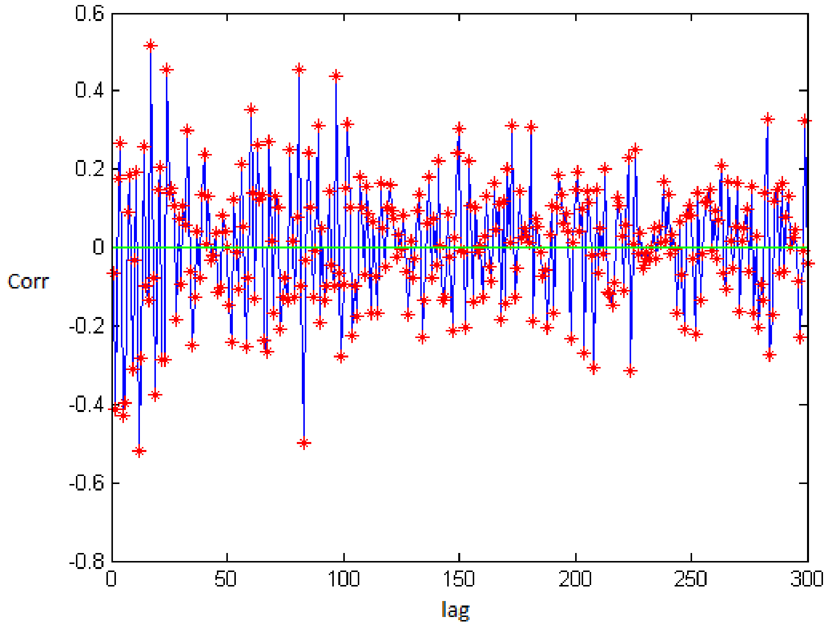

Figure 21: The correlation with the news ϵ and the price change lag p days after ΔP_{t+p} for SXX.

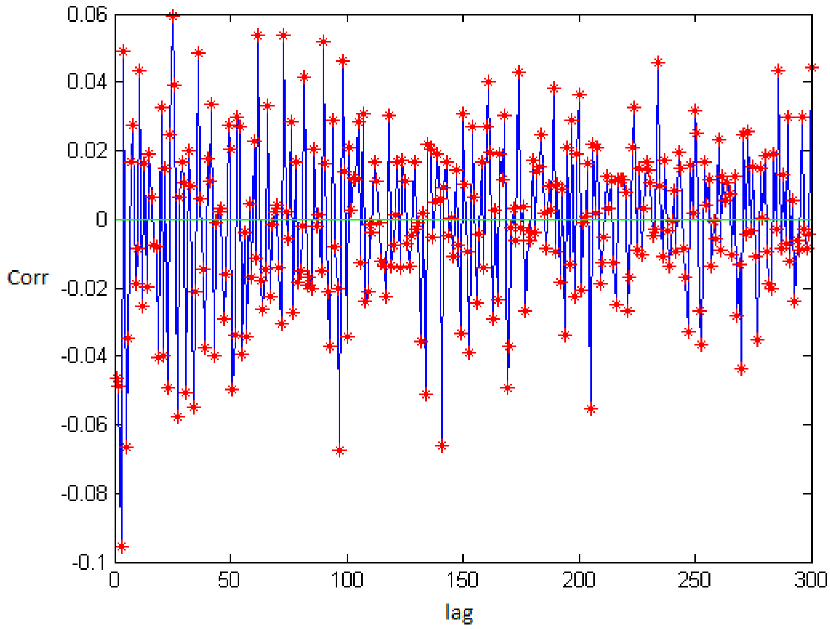

Figure 22: The correlation with the news ϵ and the price change lag p days after ΔP_{t+p} for UKX.

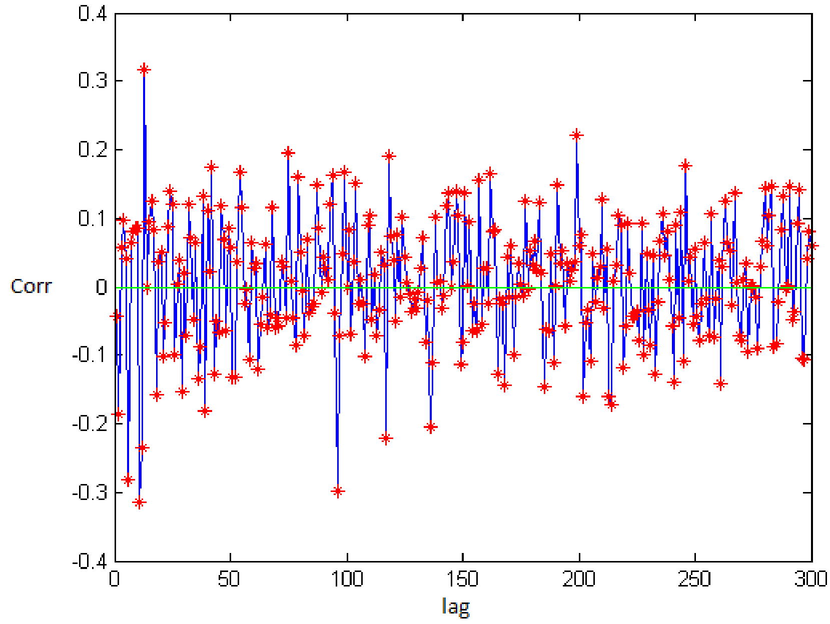

Figure 23: The correlation with the news ϵ and the price change lag p days after ΔP_{t+p} for TPX.

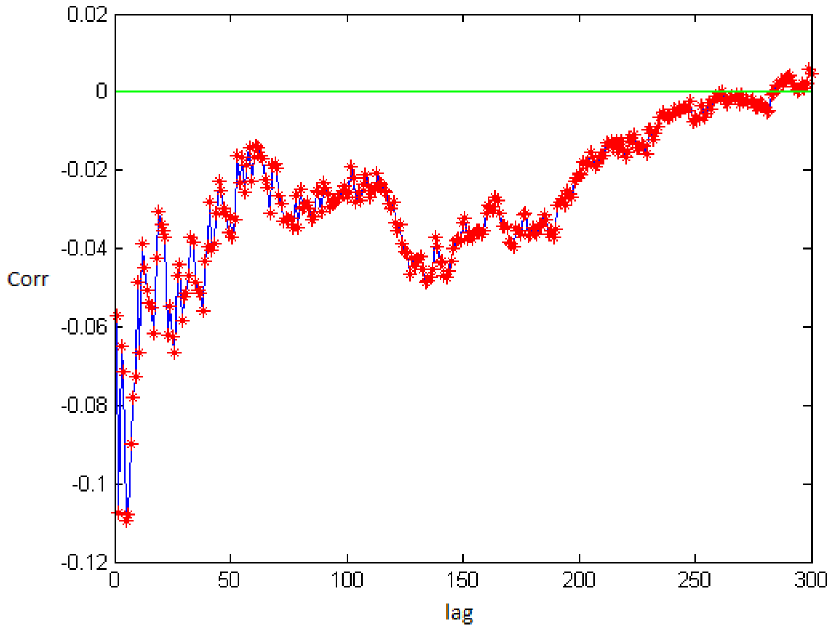

Figure 24: The correlation with the news ϵ and the price change from the day after until p days after $P_{t+p} - P_t$ for SPX.

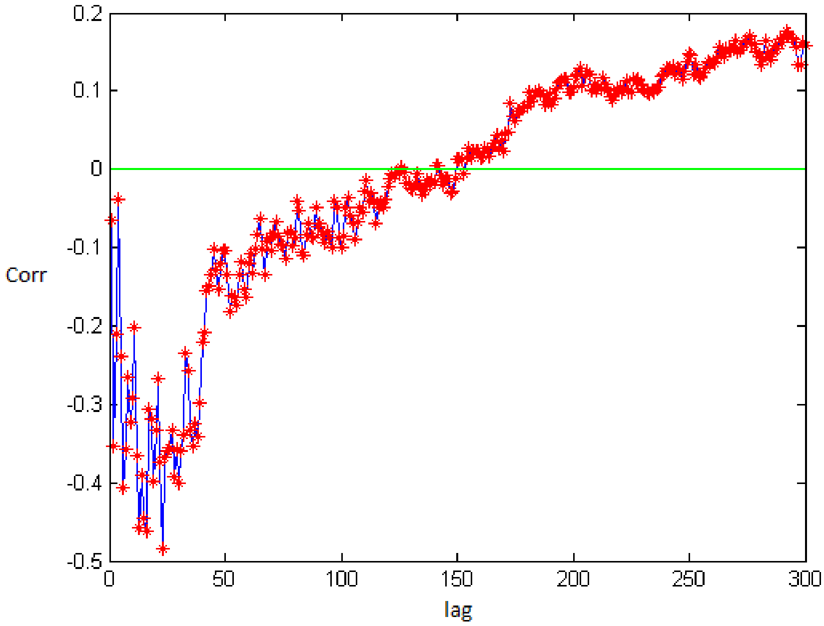

Figure 25: The correlation with the news ϵ and the price change from the day after until p days after $P_{t+p} - P_t$ for SXX.

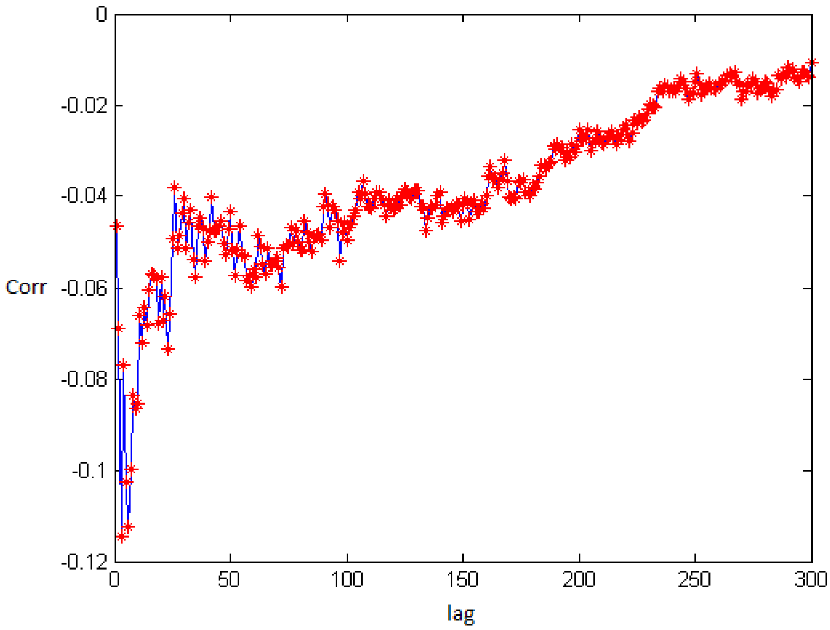

Figure 26: The correlation with the news ϵ and the price change from the day after until p days after $P_{t+p} - P_t$ for UKX.

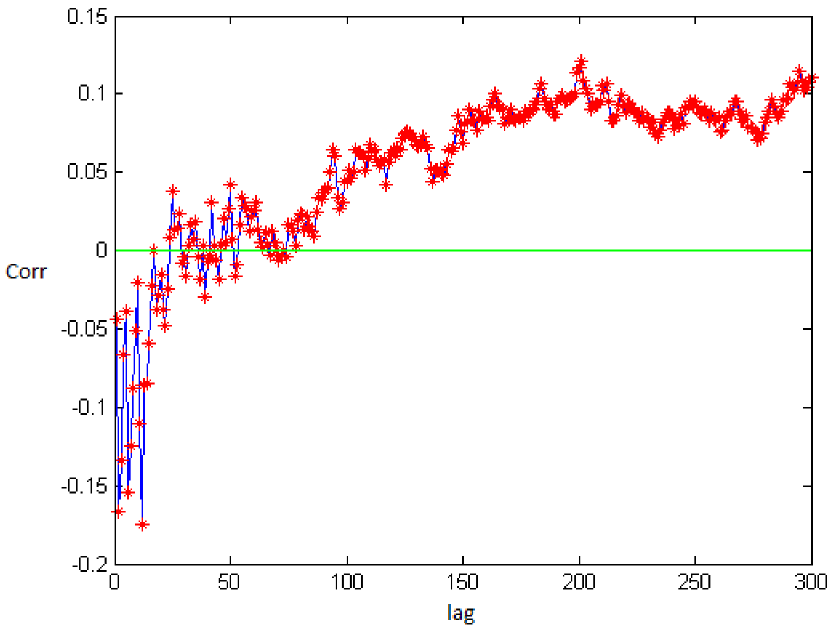

Figure 27: The correlation with the news ϵ and the price change from the day after until p days after $P_{t+p} - P_t$ for TPX.

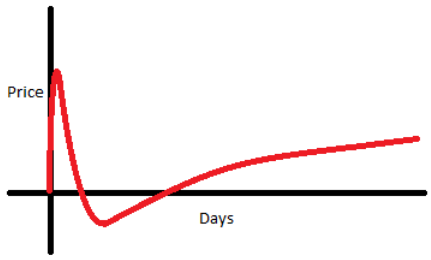

Figure 28: One interpretation of how the price is reacting on a positive news announcement at day zero.

4.3 Stocks

We randomly choose 40 stocks from the S&P 500. One group is the 20 first stocks whose name starts with P, the other group is the first 20 stocks starting with T. The information on their prices is retrieved from Google Finance[1]. The dates of their earnings announcements for the last ten years are retrieved from The Street[2].

4.3.1 Autocorrelation and partial autocorrelation functions

In an attempt to see if there are any similarities with the stylized example, we plot the autocorrelation and the partial autocorrelation functions in Figure 29 - 30 in two different cases. In both cases we find no similarities, thus our original model fails when it comes to stocks as well.

4.3.2 Correlation with the news

Next we calculate the correlation between the news (i.e. the return on the announcement day) and the return over the next p days,

$$Corr(\epsilon_t, P_{t+p} - P_t)$$

The correlation plots (Figure 33-34) do not show the pattern we saw in the indices. We will look at different explanations for this result.

[1] http://www.google.com/finance
[2] http://www.thestreet.com

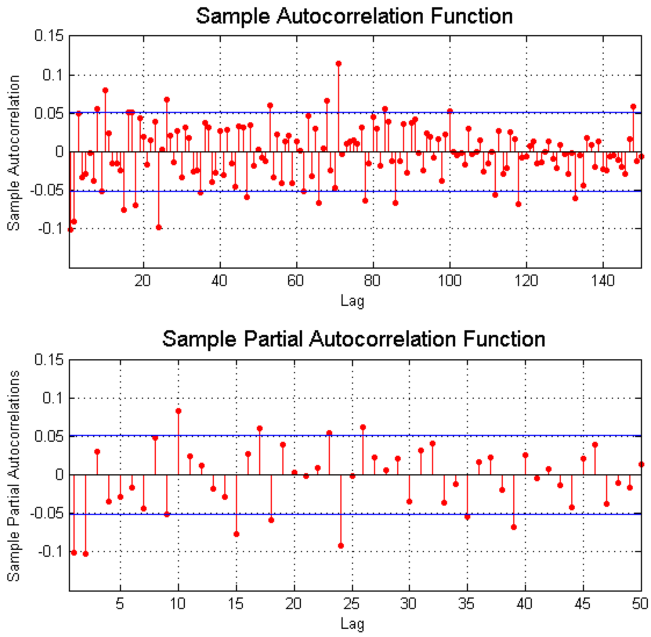

Figure 29: The autocorrelation and the partial autocorrelation function of the return from having one long position in all the P stocks.

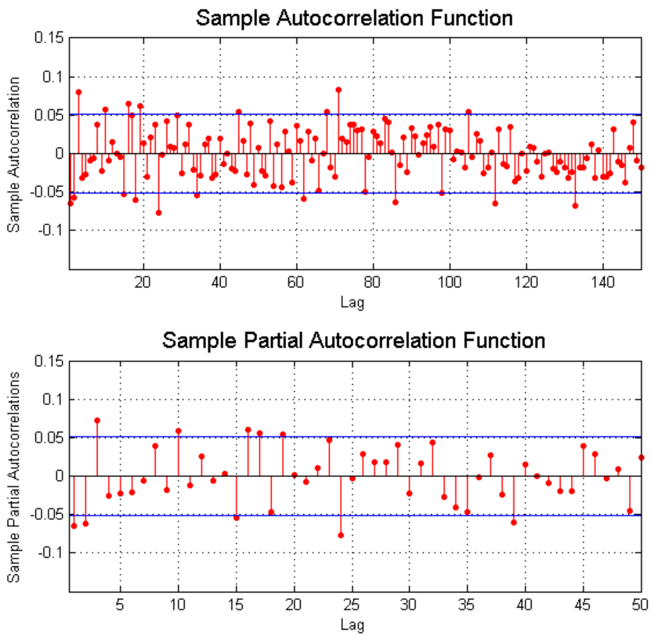

Figure 30: The autocorrelation and the partial autocorrelation function of the return from having one long position in all the T stocks.

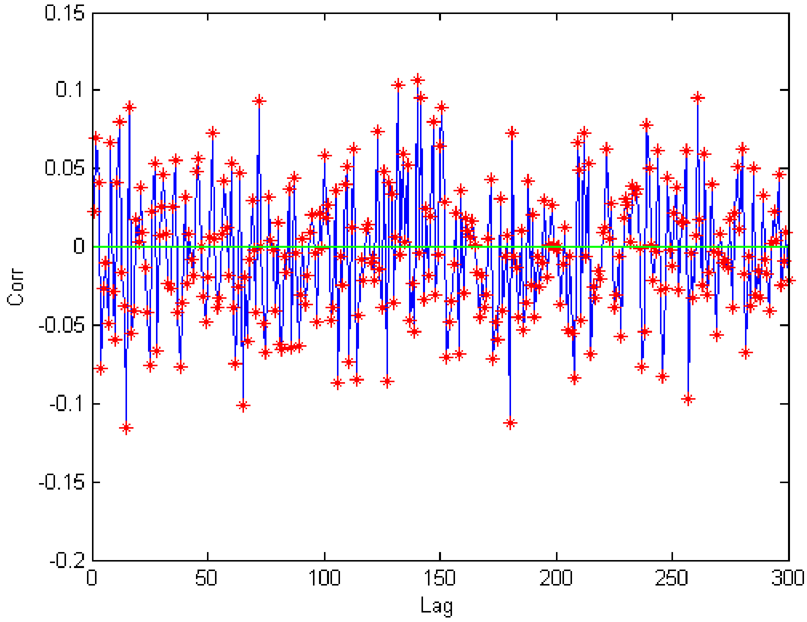

Figure 31: The correlation with the news ϵ and the price change lag p days after ΔP_{t+p} for the stocks that begins P.

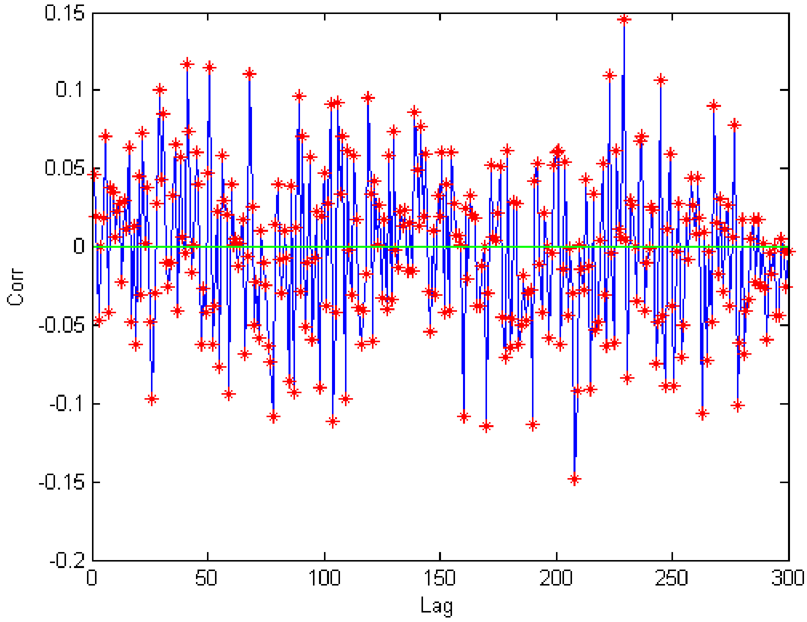

Figure 32: The correlation with the news ϵ and the price change lag p days after ΔP_{t+p} for the stocks that begins T.

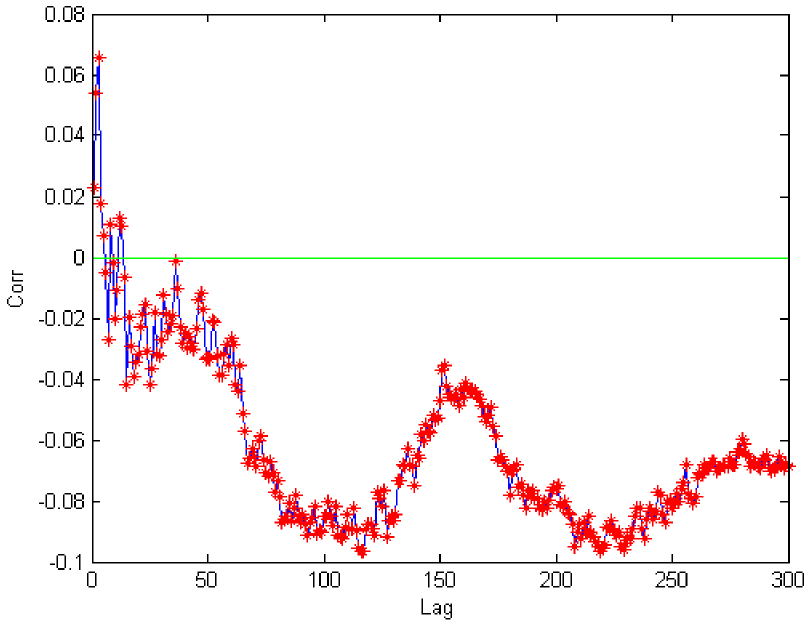

Figure 33: The correlation with the news ϵ and the price change from the day after until p days after $P_{t+p} - P_t$ for the P stocks.

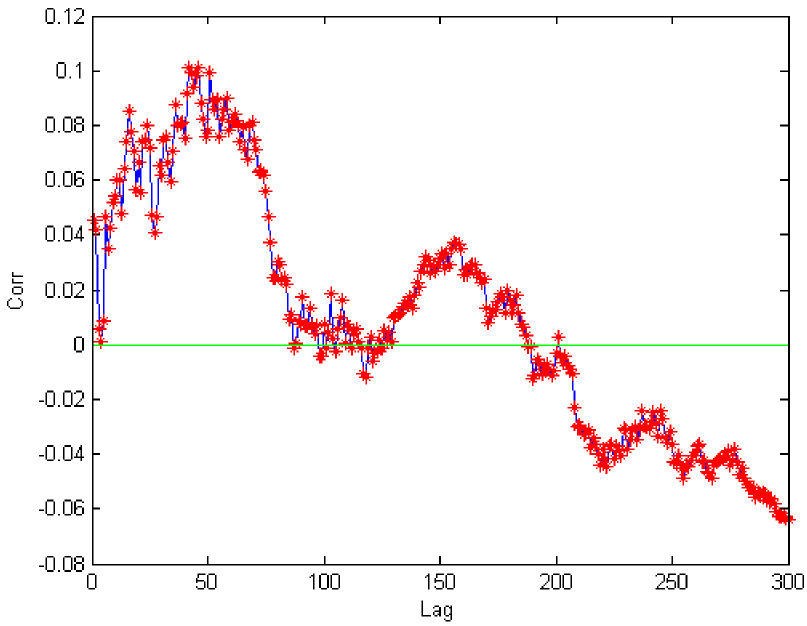

Figure 34: The correlation with the news ϵ and the price change from the day after until p days after $P_{t+p} - P_t$ for the T stocks.

4.4 First stock to announce in each sector

A possible explanation is that companies react to the earnings announcements of other companies, especially those in the same sector. To isolate this effect, we pick the companies who announced their earnings first in each sector the last quarter of 2010 and assume that these companies were the first to announce in their sector over the past ten years. A list of these companies and the corresponding sectors are presented in Table 2. The correlation between the news (i.e. the return on the announcement day) and the return over the next p days,

$$Corr(\epsilon_t, P_{t+p} - P_t)$$

for the companies who announce their earnings first is presented in Figure 35. We see the same pattern as in the indices. This evidence supports the case for the price patterns described in section 4.2.2.

Sector	Company to announce earnings of Q4 2010 first
Oil and Gas	Schlumberger Ltd
Basic Materials	Alcoa Inc
Industrials	Fastenal Co
Consumer Goods	Lennar Corp
Health Care	UnitedHealth Group Inc
Consumer Service	eBay Inc
Telecommunication	Verizon Communication
Utilities	Consolidated Edison
Financial	JPMorgan Chase & Co
Technology	Intel Corp

Table 2: The Company in each sector who announce their earnings from Q4 2010 first

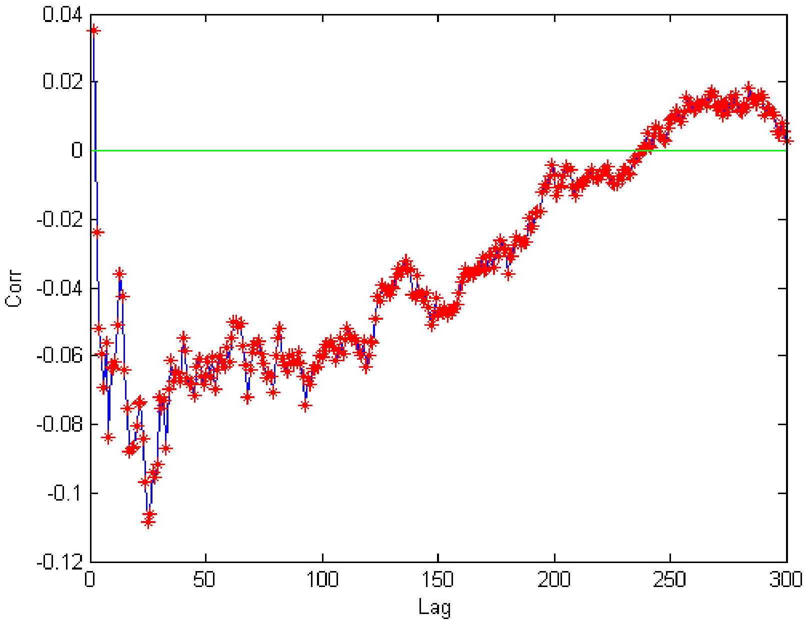

Figure 35: The correlation with the news ϵ and the price change from the day after until p days after $P_{t+p} - P_t$ for the companies who announce their earnings first.

4.5 Sectors

The last section indicates that there is a significant correlation amongst companies in the same sector. Let us see if the first company to announce can forecast the sector behaviour. We use the sector iShares as a proxy for sector performance (Table 3). Define ϵ_t as the price change of the index on the day that the first stock in the sector announces its earnings. We then plot in Figure 36 the correlation between ϵ_t and the price change of the sector from the day after the first company's announcement until p days later $P_{t+p} - P_t$. The correlation plot does not show the same clear patterns we observed for stock market indices and the first stocks in the sector. Therefore the information coming from the first company cannot be seen as representative for the entire sector. Furthermore, sectors have their earnings season spread out over the entire SPX index earnings season and thus are affected both by sector-specific and market-wide news.

Sector	Index
Oil and Gas	iShares Dow Jones US Oil & Gas Ex Index (IEO)
Basic Materials	iShares Dow Jones US Basic Materials (IYM)
Industrials	iShares Dow Jones US Industrial (IYJ)
Consumer Goods	iShares Dow Jones US Consumer Goods (IYK)
Health Care	iShares Dow Jones US Healthcare (IYH)
Consumer Service	iShares Dow Jones US Consumer Services (IYC)
Telecommunication	iShares Dow Jones US Telecom (IYZ)
Utilities	iShares Dow Jones US Utilities (IDU)
Financial	iShares Dow Jones US Financial Sector (IYF)
Technology	iShares Dow Jones US Technology (IYW)

Table 3: The index used for the different sectors.

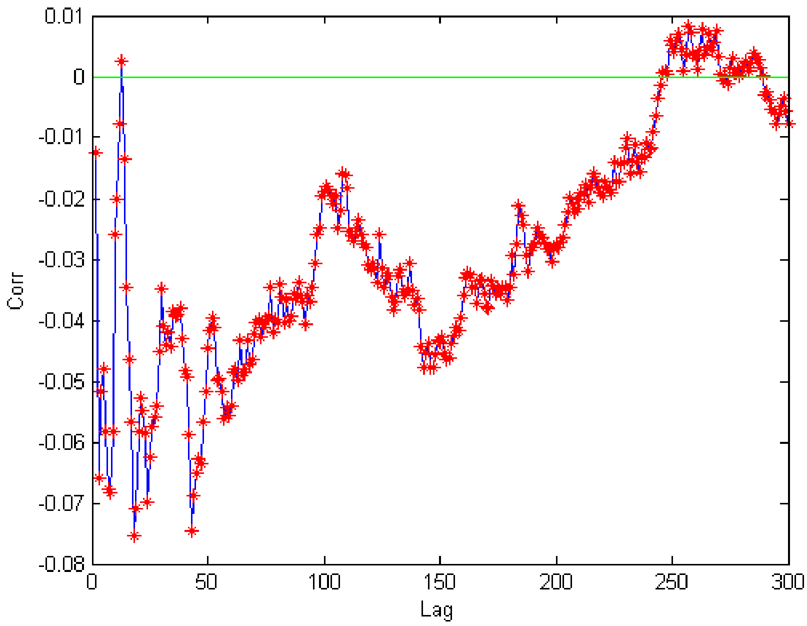

Figure 36: The correlation with the news ϵ and the price change from the day after until p days after $P_{t+p} - P_t$ for the sector indexes.

5 Test of a simple trading strategy

The plots of the autocorrelation and the partial autocorrelation function Chapter 4 suggest that the market reaction to earnings news differs from that predicted by the model. There is an overreaction on the day of the announcement, a correction that lasts for 5-10 days and overshoots the price in the opposite direction and eventually a long trend in the same direction as the price move on the day of the announcement. To test the validity of this observation depicted in Figure 28, we devise a trading strategy that exploits this behaviour and test its performance. To measure performance we use the information ratio (IR) defined as:

$$IR = \frac{E[r_{annual}]}{\sqrt{Var(r_{annual})}}$$

At first the same four indexes used in the last chapter will be used and the strategy will be developed step by step. We will then test the strategy out of sample by applying it to four new indices. These indices are OMX (Sweden), AS51 (Australia), SMI (Swiss) and DAX (Germany). Finally, we will see how well the strategy works on the stocks that had their earnings announcements first in each sector.

5.1 Equity futures

Although the benchmark for the strategy is zero, i.e. the strategy is an alpha strategy, it is interesting to see what happens if we always hold one long position in all indices, see Figure 37. The IR for the long-only strategy is 0.14.

Then we test the following strategy: wait five days after an earnings announcement, then take a position in the same direction as the market moved on the day of the announcement and hold it for a year. This means that during the earnings season, we take a new position every day in the direction the market moved five days ago. The NAV of the this strategy is shown in Figure 38. The IR value is 0.35. We include a trading cost of 0.1%.

The price reaction sketched in Figure 28, suggests that it would be profitable to take a position for the first four days after the announcement in the opposite direction of the price change on the day of the announcement. This strategy is shown in Figure 39 and the IR value has now increased to 0.46 (it reaches 0.64 if only the last five years are considered).

One could believe that the difference between the last strategy and the long-only is due to the profit stemming from chasing a trend, which is unrelated to the earnings announcements. To check this we implement a simple trend chasing strategy: take a position in the same direction that the market moved yesterday and hold for a year. The performance of this strategy is shown in Figure 40 and the corresponding IR value is 0.23. So even if it beats the long-only strategy, it is relatively far away from the IR values connected with the strategy that builds on price behaviour during earnings seasons.

Can we improve on the strategy even more? It would make sense to only take a position if the price change on announcement day was significant. So the strategy checks if the price change on the day of the earning

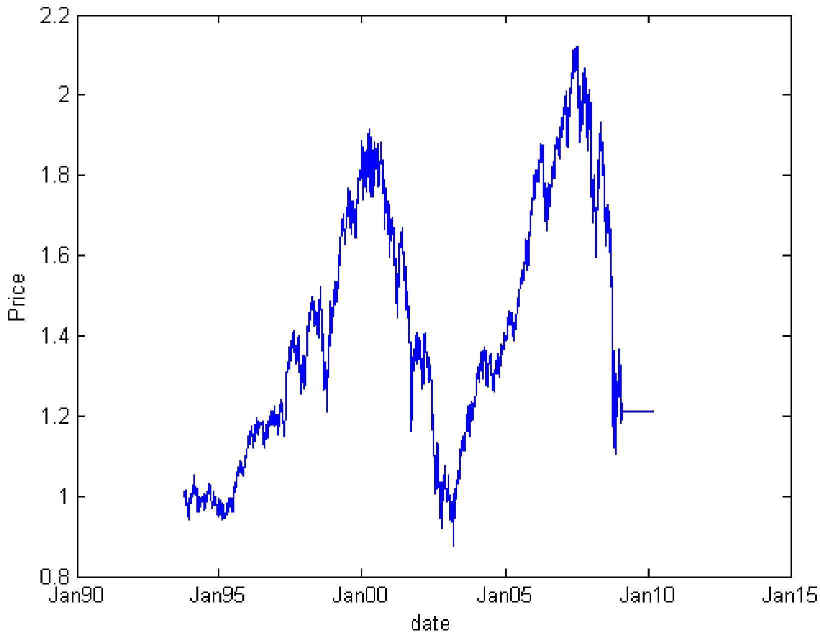

Figure 37: Taking one long position in all the four indexes.

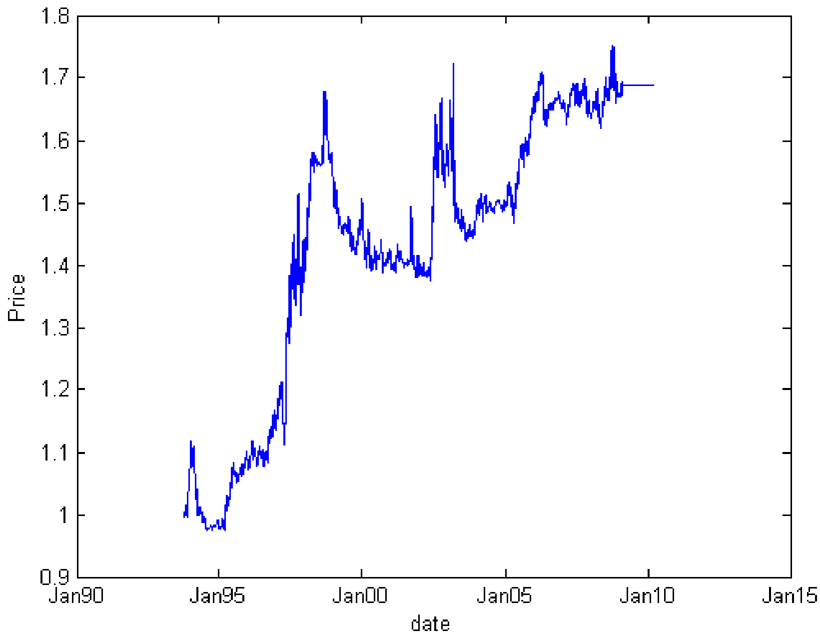

Figure 38: Wait five days after an earnings announcement then takes a position in the same direction

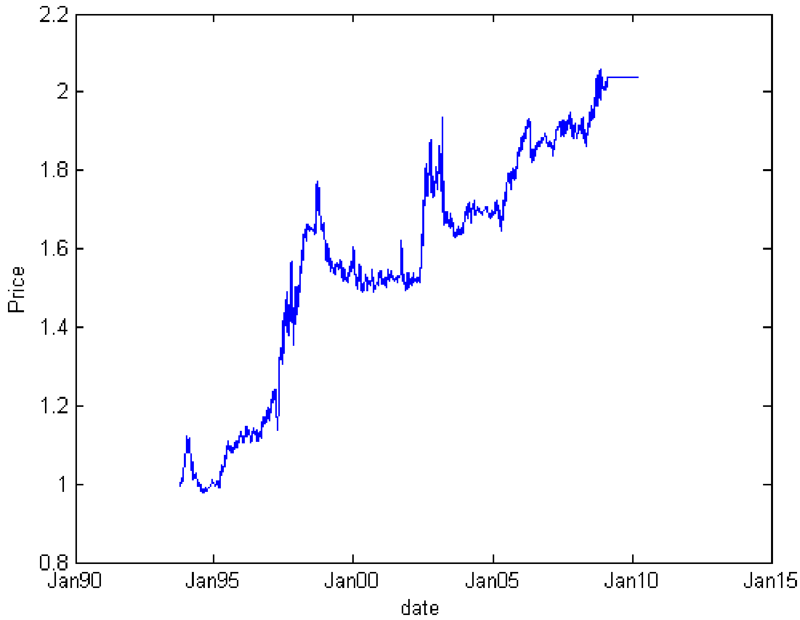

Figure 39: Take an opposite position as the earnings announcement for the first four days then do the opposite.

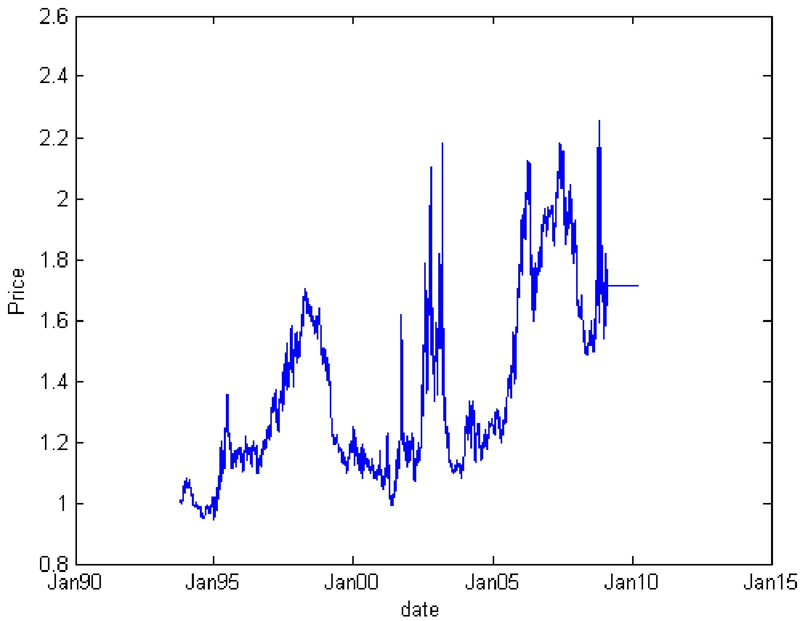

Figure 40: Follow a simple trend chasing strategy, take a position in the same direction as the price change yesterday and hold for a year

announcement is larger than one standard deviation of the returns over the last half year. If this is this is the case then takes the opposite position for the first four days followed by a position in the direction of the announcement return from day five up until a year. This strategy has an IR of 0.72 (for the last five years the IR = 1.24). It is shown in Figure 41 and with statistical data presented in Table 4. If we increase the threshold to two standard deviations of daily returns, the strategy is inactive for long periods; we do not present this result. The trading strategy is predicated on the release of earnings-related news. To gauge the impact of this, instead of basing the strategy on the earnings season we base it on what happens one or two months later (i.e. we move one month and two months forward the period over which the Dirac delta function takes the value one in the definition of $\epsilon_{t+i} = s\delta_0(i)\Delta P_{t+i-z},\ i = 0, 1, ..., z-1,)$. This gives IR values of 0.27 and 0.38 for the entire period and with just the last five years 0.43 and 0.23, both significantly below the values of the strategy tied to the actual earnings announcements.

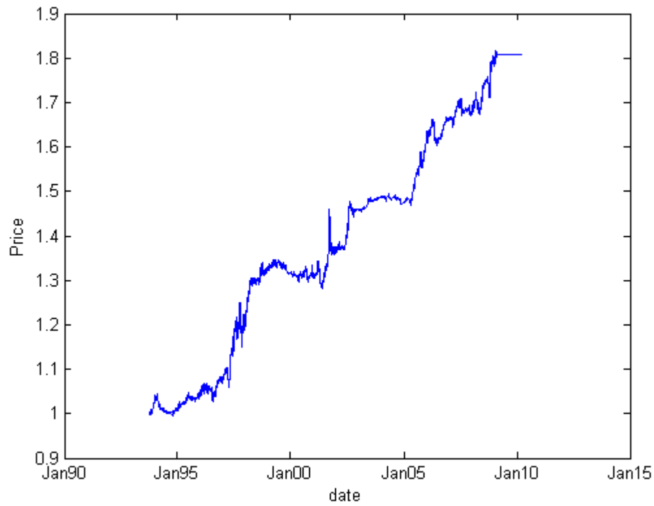

Figure 41: Take opposite positions as the earnings announcement for the first four days then do the opposite if the price change was higher than one standard deviation.

Finally we test the strategy out-of-sample. We apply it on four new indices: OMX (Sweden), AS51 (Australia), SMI (Swiss) and DAX (Germany). Historical returns are retrieved from Bloomberg, along with details of their first earnings announcement season in 2011. The earnings season for each market is determined by looking at the concentration of earnings announcements, these seasons are presented in Table 5. We make the same assumption here as for the previous four indexes: these earnings seasons had been the same during the last 15 years.

The performance of the trading strategy developed is presented in Figure 42 and Table 6. The IR value has now dropped to 0.44 (0.68 the last five

Best return	0.0356
Worst return	-0.0359
Skewness of return	0.21
Excess kurtosis of return	20.17
Percentage of up days	0.49
Percentage of down days	0.44
Information ratio (IR)	0.72
IR last three years	1.06
IR last five years	1.24
IR last ten years	0.95
Draw-down to volatility	-1.95

Table 4: A table with statistical data of the return related to Figure 41

Index	Start date	End date
OMX	Fourth Monday of January	Second Friday of February
AS51	First Monday of February	First Friday of March
SMI	First Monday of February	First Friday of March
DAX	Third Monday of February	Third Friday of March

Table 5: The period of when the earnings announcements are for the different markets for the last quarter of 2010.

years)

If we move the period, over which the Dirac delta function in ϵ takes the value one, one and two months forward, we get IR values of 0.08 and -0.13 for the entire period and with just the last five years 0.17 and -0.15, both significantly below the values based on the true earnings season. So even though the IR values are lower, the degradation in IR when we debase the strategy from earnings is about the same.

Best return	0.0217
Worst return	-0.0380
Skewness of return	-0.64
Excess kurtosis of return	27.19
Percentage of up days	0.44
Percentage of down days	0.42
Information ratio (IR)	0.44
IR last three years	0.57
IR last five years	0.68
IR last ten years	0.30
Draw-down to volatility	-3.70

Table 6: A table with statistical data of the return related to Figure 42

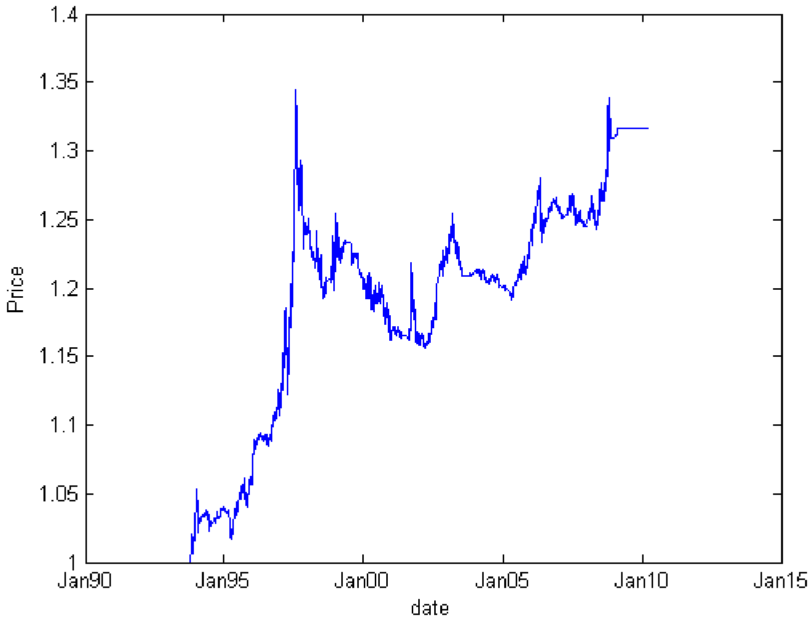

Figure 42: Take opposite positions as the earnings announcement for the first four days then do the opposite if the price change was higher than one standard deviation.

5.2 Stocks

We apply the above trading strategy to the ten stocks that announce first in the ten sectors: if the price change on the day of the announcement is larger than one standard deviation of returns, we take a position in the opposite direction for the following four days, then reverse the position and hold it for a year. Every time we trade, we incur a trading cost of 0.1%. The result is presented in Figure 43 and in Table 7. The IR is 0.64; this can be compared to an IR value of 0.17 for the long-only strategy. It is also noted that if instead of the date of the earnings announcement we use the day after or the day before, the IR value drops to -0.44 and -0.51 respectively. Thus the importance of taking the exact earnings release dates is great.

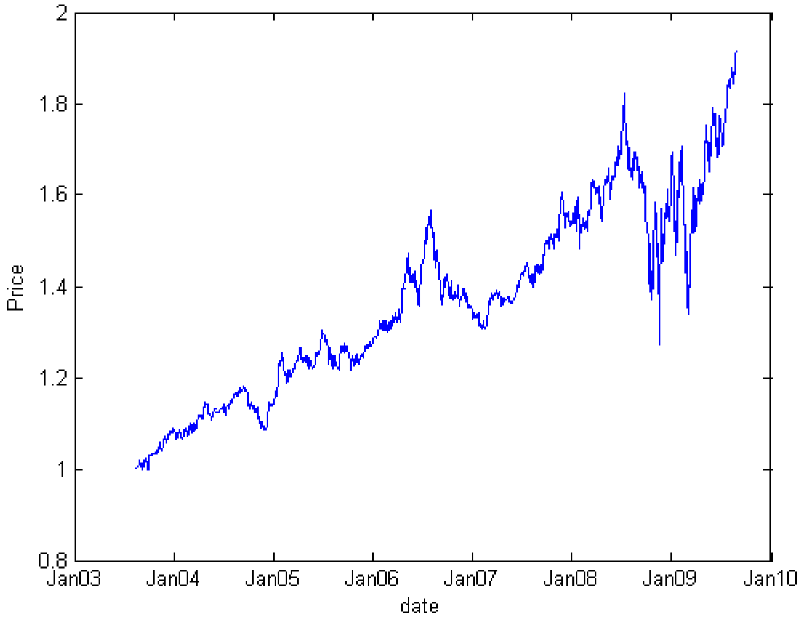

Figure 43: The trading strategy on the 10 stocks

Best return	0.0784
Worst return	-0.0706
Skewness of return	0.23
Excess kurtosis of return	10.50
Percentage of up days	0.54
Percentage of down days	0.46
Information ratio (IR)	0.64
IR last three years	0.40
IR last five years	0.54
Draw-down to volatility	-1.71

Table 7: A table with statistical data of the return related to Figure 43

6 Additional tests of significance

We use three different tests to investigate the significance of the returns given by the trading strategy. At first we will look at t-statistic based on an article by Donoho, Crenian and Scanlan (2010), where they argue that even a bad strategy can have a good return in the short run. Therefore, the time of which the strategy has been applied on is of great importance. Secondly, we test the null hypothesis that the returns are independently identically distributed (IID) and identical, both with and without the strategy. If the strategy is informative, is should alter the distribution. This is done with two different goodness of fit tests, by comparing the quantiles and a Kolmogorov-Smirnov test. Finally, we bootstrap the returns and run several samples of the Kolmogorov-Smirnov test.

6.1 t-statistics

If we are considering the actual return generated over the last Y years, adjusted for the market exposure as our only information. The t-statistics of our strategy performance is given by (Donoho, Crenian and Scanlan (2010));

$$t_s := \frac{E[r_{annual}]}{\sqrt{Var(r_{annual})}}\sqrt{Y} = IR\sqrt{Y},$$

The IR is the same IR calculated in our trading strategy for the returns. This yields Table 8, the convention of when a t-statistic number implies statistical significance varies from field to field. Donoho, Crenian and Scanlan (2010) set the limit of significance at 2.

Indices/Stocks	t_s last five years	t_s entire period
Indices (SPX, UKX, TPX, SXXE)	2.8	3.0
Indices (OMX, AS51, SMI, DAX)	1.5	1.8
Stocks	1.2	1.7

Table 8: t-statistics of the returns

6.2 Goodness of Fit

If the two normalized return distributions, the one with and the other one without the strategy are identical their quantiles should be close to identical (Lo, Mamaysky, Wang (2000)). To check this, we compute the deciles of normalized returns without the strategy and tabulate the relative frequency δ_j of normalized returns with the strategy falling into decile j of the normalized returns without the strategy.

$$\delta_j \equiv \frac{\text{number for normalized returns with the strategy in decile j}}{\text{total number of returns with the strategy}},$$

We normalize the returns by subtracting means and dividing by the standard deviation, hence;

$$r_{normalized} = \frac{r - mean(r)}{std(r)},$$

Under the null hypothesis that the returns are IID and the returns with and without the strategy are identical, the corresponding goodness of fit test statistic Q are given by (Lo, Mamaysky, Wang (2000));

$$\sqrt{n}(\delta_j - 0.1) \overset{d}{\sim} N(0, 0.1(1-0.1)),$$

$$Q \equiv \sum_{j=1}^{10} \frac{(n_j - 0.1n)^2}{0.1n} \overset{d}{\sim} \chi_9^2,$$

where n_j is the number of observation that fall in decile j and n is the total number of observation. The δ_j and Q values for our strategy is tabulated in Table 9. The null hypothesis is rejected for both the indices and the stocks at significance level 0.01.

Indices/Stocks	SPX,UKX,TPX,SXXE	OMX,AS51,SMI,DAX	Stocks
Decile 1	0.079	0.065	0.089
Decile 2	0.070	0.065	0.102
Decile 3	0.095	0.089	0.100
Decile 4	0.140	0.113	0.103
Decile 5	0.162	0.249	0.114
Decile 6	0.123	0.135	0.101
Decile 7	0.093	0.082	0.120
Decile 8	0.077	0.071	0.074
Decile 9	0.070	0.059	0.084
Decile 10	0.089	0.069	0.107
Q	362.5	1197.2	26.5
p-value	0.000	0.000	0.002

Table 9: Goodness of fit with quantiles

6.3 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test is based on the empirical distribution function $F_i(x)$, which for n IID observations X_{ik} is defined as;

$$F_i(x) \equiv \frac{1}{n_i} \sum_{k=1}^{n_i} I_{X_{ik} \leq x}$$

where $I_{X_{ik} \leq x}$ is the indication function, equal to 1 if $X_{ik} \leq x$ otherwise 0. The test is designed to test the null hypothesis that $F_1 = F_2$. The statistics is given by the expression;

$$\gamma_{n_1,n_2} = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \sup_x |F_1(x) - F_2(x)|$$

Under the null hypothesis, $F_1 = F_2$, γ_{n_1,n_2} is zero. In particular, the hypothesis cannot be rejected at level α if $\gamma_{n_1,n_2} \leq K_\alpha$, where K_α is found from $Pr(K \leq K_\alpha) = 1 - \alpha$. K is Kolmogorov distributed;

$$Pr(K \leq x) = 1 - 2\sum_{i=1}^{\infty}(-1)^{i-1}e^{-2i^2x^2}$$

With our normalized returns, the null hypothesis is rejected for both groups of indices at significance level 0.01, however, the null hypothesis for the returns of the stocks cannot be rejected at significance level 0.05.

We use the n number of returns with the strategy to bootstrap (with replacement) 1000 new samples, each of them with n number of returns. We run the Kolmogorov-Smirnov test with the same null hypothesis with all these samples. For both groups of indices all samples are rejected at significance level 0.01. With the stocks 151 of the samples are rejected at significance level 0.05 and 37 samples are rejected at significance level 0.01.

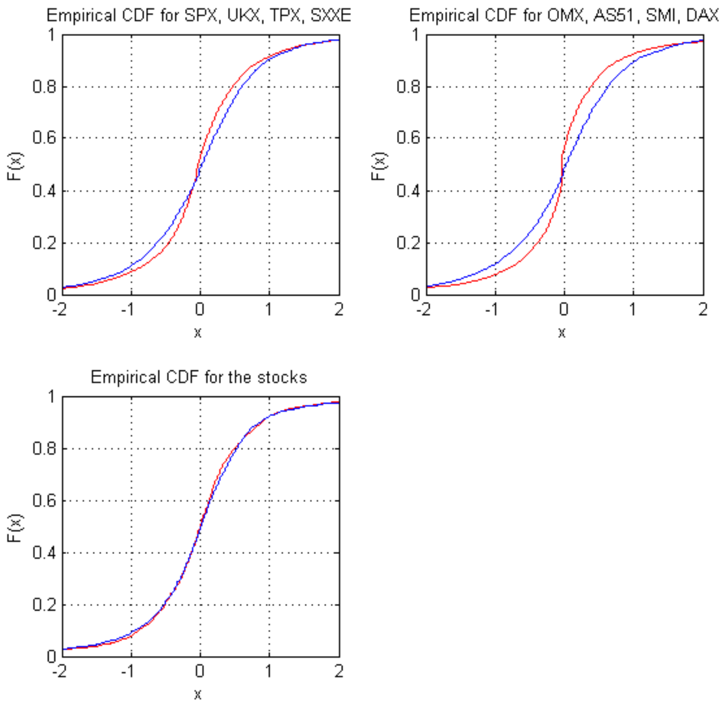

Figure 44: The empirical distribution functions. The red line is the returns with the strategy, the blue line is the returns without the strategy.

7 Conclusion

Hong and Stein's (1999) model is based on three main assumptions: (a) that the market only consists of two types of traders, news-watchers and momentum traders, (b) that private news gradually diffuses across the population and (c) that momentum traders use a simple strategy only conditioned on yesterday's price change. If this were the case, the model's parameters would have been identifiable in the correlation plots. However, they are not identifiable in the correlation plots and therefore the model cannot explain the complex reality.

As Hong and Stein (1999) pointed out, if the news is public, smart momentum traders would also condition on the release dates thus responding quicker to the news. This can explain what we see with the indices and with the stocks who are the first ones to announce in each sector.

So can we add this additional momentum trader to the model? The answer is no because our news definition is the market return on the day of the announcement. We would need to sample prices at a higher frequency during the announcement day (e.g every hour) and use these price returns as the news fed to the model.

Another explanation of why the stocks seem to react a lot quicker than the model is describing is that the stocks used here are all very liquid. This means that they are easily traded, thus information can be incorporated in the price quicker than for illiquid stocks. Chordia et.al. (2007) shows that post earnings announcements drift are not that severe on liquid markets as on illiquid markets.

Our findings indicate that there is some post drift from the earnings announcements and it seems to be an overreaction on the announcement day. The trading strategy built around this observation easily beats the long-only and a simple trend chasing strategy and loses all its power when it is detached from the announcement timetable. This is in accordance with findings of Frazzini and Lamont (2006), who show that stock returns are on average abnormally high during the day of the earnings announcements.

During periods of news announcements, the trading volume goes up, see Hong and Stein (2006). This can explain why Piqueira (2006) finds a significant negative association between turnover and future returns. This supports the strategy used in Chapter 5.

Why the model on indices works better for the last five years, can be explained by the fact that we took the first earnings season of 2011 and assumed it had been the same every quarter of the past 15 years. Unfortunately, we have been unable to find detailed data for past earnings seasons.

We showed that earnings information is important for stocks that report first in their sector. For these stocks we found post-earnings patterns similar to those of market indices. The importance of earnings for a random stock seems to be limited.

References

[1] Hong, H., J.C. Stein, 1999, A unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, *Journal of Finance*, 54, 2143-2184.

[2] Jegadeesh, N., S. Titman, 1993, Returns to buying winners and selling losers; Implications for stock market efficiency, *Journal of Finance*, 48, 65-91.

[3] Bernard, V. L., 1992, Stock price reactions to earnings announcements, in Richard Thaler, ed.: *Advances in Behavioral Financial* (Russel Sage Foundation, New York)

[4] DeBondt, W. F. M., R. H. Thaler, 1985, Does the stock overreact?, *Journal of Finance*, 40, 793-808.

[5] Schwert, G. W., 2003, Anomalies and market efficiency, in G.M. Constantinides M. Harris R. M. Stulz (ed.), *Handbook of the Economics of Finance*, edition 1, volume 1, chapter 15, 939-974.

[6] Chordia, T., A. Goyal, G. Sadka, R. Sadka, L. Shivakumar, 2007, Liquidity and the Post-Earnings-Announcements-Drift, *AFA 2008 New Orleans Meetings Paper*, 1-34.

[7] Lo, A.W., H. Mamaysky, J. Wang, 2000, Foundation of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, *Journal of Finance*, 55, 1705-1765.

[8] Donoho, D.L, R. A. Crenian, M. H. Scanlan, 2010, Is Patience a Virtue? The Unsentimental Case for the Long View in Evaluating Return, *Journal of Portfolio Management*, 105-120.

[9] Frazzini, A., O. Lamont, 2006, The earnings announcement premium and trading volume, Yale University Working Paper.

[10] Piqueira, N., 2006, Trading activity, illiquidity costs and stock returns, University of Houston Working Paper.

[11] Hong, H., J.C. Stein, 2006, Disagreement and the stock market.

[12] Brockwell, P.J., R. A. Davis, 1991, *Time Series: Theory and Methods*, 2nd ed. Springer.

[13] Franke, J., W. K. Härdle, C. M. Hafner, 2004, *Statistics of Financial Markets: An Introduction*, 1st ed. Springer.