News analytics; analyzing the relationship between news sentiment and gold prices

By Nikki Be
Supervisor: Dr. S.A. Borovkova

Purpose

The purpose of this paper is to analyze the dependencies between the news and the price of commodities and ultimately construct a strategy that can produce a profit just by analyzing the news. As we all know, prices of assets and commodities are driven by the sentiment of the investors. This sentiment is mainly driven by the news. Negative news can really drive asset prices down and vice versa. Now Thomson Reuters has quantified this sentiment with an engine that can read and interpret news articles and give sentiment scores to each article, given the asset or commodity at matter.
In an ideal case we would find an algorithm that uses this news engine to predict movements in the price in such a way that there is a possibility to make a profit.
A greater purpose of this paper would be to analyze whether news sentiment, and the way it is processed by Thomson Reuters, would be able to give a good indication of future movements in the market. In other words, we would like to analyze the prediction power of the TRNE (Thomson Reuters News Engine).
The TRNE is a data collection of all the news articles with sentiment scores for the assets and commodities concerned. These sentiment scores are generated by an algorithm developed by Thomson Reuters and will give an indication of the sentimental value of an article towards the asset concerned.
This paper is focused around the commodity gold, as gold has been in the news a lot lately. This is mainly because of the rising gold prices of the last years, even in the most recent financial crisis. And the gold price has been rising ever since; it has even reached record high levels. Interesting to see would be whether the news has a significant influence on the gold price even on an intraday (high frequency) basis. Could negative news be the end of this Gold rush? Or is it the other way around and is it the news that becomes negative when the price drops?
When looking at the gold news we see many articles with titles like: 'Gold Price Rockets as Stock Markets Suffer', 'Biggest Bloodbath since Financial Crisis' or 'Gold Price Comes Close to Record Closing High'. These articles are not indicating sentiment for the future gold price, but are just evaluating the movement of the gold price in the past. Of course these articles do influence investor's sentiment.
The main question therefore is: does the TRNE have any prediction power, and maybe even be able to produce a profit with the right strategy based solely on the sentiment scores produced by this engine?

Table of contents

Inhoudsopgave

Purpose 2
Table of contents 3
Paper overview 4
Thomson Reuters News Analytics 5
Dataset 5
Gold 6
Gold Fixings 6
Futures 7
Aggregation 8
Aggregation methods 9
1. Mean 9
2. Bounded 9
3. Weighted 9
4. Weighted+bounded 11
Rolling Sentiment 11
Correlations 13
Daily correlations 13
High frequency correlations 14
Event studies 14
Event dates 15
Event studies on daily data 17
Event studies on intraday data 19
Investment Strategy 20
Strategy 21
Setup 21
Results 21
Conclusion 23
Future research 24
Appendix A 25
Appendix B 26
References 27

Paper overview

This paper studies the influence of news to prices of commodities, gold in particular. When clear influences can be found, an engine can be build to predict movements of the gold price caused by the news. One issue at hand is the question whether news influences the price or the price influences the news. Instinctively both will be the case, when the price is very volatile, a lot of news will be published about these movements. And also when there is a lot of news about certain commodities, the prices of these commodities will move accordingly.

The first part of this paper will be about the data that will be used to analyze the relation between the value of gold and the news sentiment. It will give an insight in the dataset that Thomson Reuters has provided, with the news sentiment scores. As well as the gold prices that will be used. Throughout this research both daily gold prices and high frequency (intraday) gold prices will be used.

When having a clear insight in the datasets that will be used, the data processing will be explained. As with all raw data, it first has to be processed to be usable for further research. The sentiment scores provided by Thomson Reuters have to be aggregated in order to be usable for further research. When all data is processed the search for dependencies between gold and the news can begin.

From the aggregation, a rolling sentiment score can be created to find a score for the sentiment at each point in time. This score then will be compared to the changes in gold price (log-returns), by means of correlation coefficients.

Also, so-called event studies to see what happens around extreme news events. Extreme news events are scores that are in the upper quantile of the dataset. It is very interesting to see what happens with the sentiment and the cumulative returns around these events.

Finally, an investment strategy will be implemented to see to what extend the news can predict movements in gold price and if there are any profits that can be made from this information.

Thomson Reuters News Analytics

Due to growing interest in investing in commodities, Thomson Reuters has extended their News Analytics engine in 2009. Now, sentiment scores are also available for articles related to several commodities, such as natural gas, crude oil and gold. This news sentiment engine generates a sentiment value to the positivity/negativity of the article regarding the commodity at matter. These values represent the probabilities of the article to be positive, neutral or negative. The values of the positive, neutral and negative sentiment therefore always add up to 1. Unfortunately, the way the scores are produced is black box, i.e. I don't have any insight in how the scores are calculated. Therefore I can only assume these scores are representative for the article, and truly represent the news sentiment of an article.

Dataset

The Thomson Reuters news engine provides a dataset with 82 variables, including the sentiment scores. The dataset includes articles, appends (extensions of articles), alerts (small news flashes) and rewrites (copies of articles). I will only use the articles in this paper.
In the first phase of the chapter I will only use the data that is shown in Table 1.

IDN_TIME	REL	SENT	SENT_POS	SENT_NEUT	SENT_NEG
37622.67633	1	-1	0.230784	0.268766	0.50045
37623.07476	1	-1	0.231234	0.323681	0.445085
37623.13449	0.059	1	0.430833	0.394884	0.174283
37623.23709	1	1	0.610747	0.106723	0.28253

Table 1: small sample of the news data

Where IDN_TIME gives the excel number for a specific point in time (GMT). REL is the relevance of the article; all articles with a relevance smaller than 0.3 will be considered irrelevant and will be removed from the dataset. The 'SENT' gives the sentiment (1 for positive, -1 for negative, 0 for neutral). Also the probabilities for these sentiments are shown.

The news data, ranging from January 2003 until September 2011, contains 273412 articles related to gold, ranging from around 1500 up to over 4000 articles per month.

In Figure 1 the distribution of the articles over the week is shown in blue. We can see a regular distribution of news throughout the week; more news in the middle of the week and less news in the weekend. The red columns represent the number of articles that influence the price on the corresponding day of the week. When there are no returns, the articles of that day will be aggregated with the articles of the next day, i.e. the red bar on Monday is the total number of articles between Friday 10:30AM and Monday 10:30AM (as there are no returns in the weekend), this is the publishing time of the gold fixings (later more on that).
A more specific distribution of the news articles is given in appendix A.

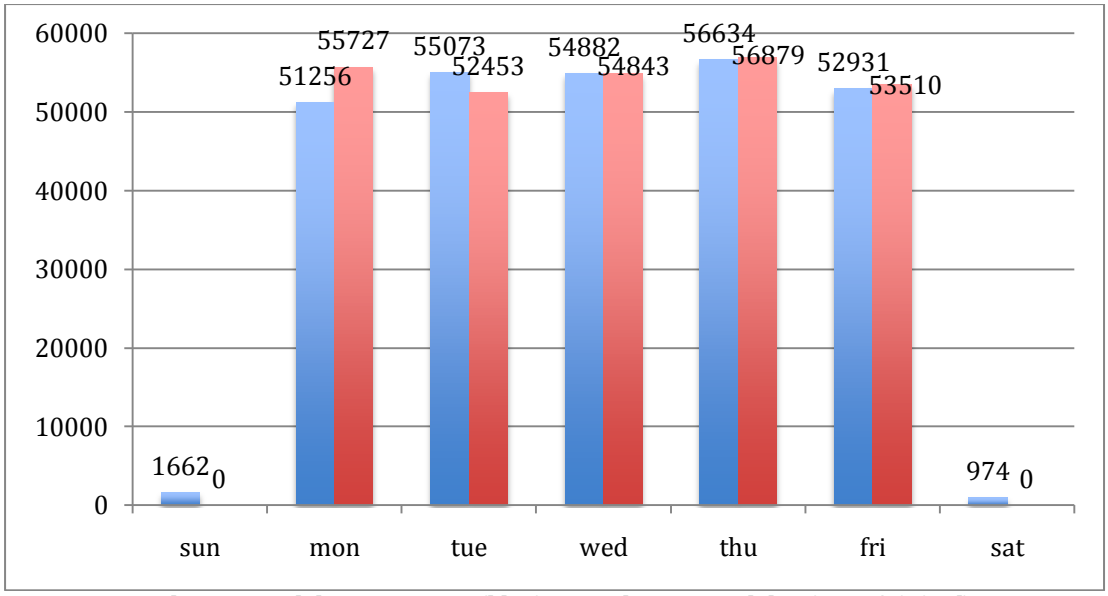


Figure 1: #articles per weekday 0:00 - 0:00 (blue), #articles per weekday (mon-fri) (red)

Gold

This paper focuses on gold, as it has been in the news very frequent lately. More and more investors are interested in investing in gold, as the value has been rising for an extended period of time now. Since the latest financial crisis, investing in gold has become particularly interesting. Many economists claim that gold is extremely stable, and especially in crisis times can be a very stable investment. In this paper, both daily and intraday data is used for researching the relationship between the news sentiment and the gold price. For the daily data I will use the gold fixings, published in London. And for the intraday data I will use the gold future bid-ask quotes, provided by Thomson Reuters.

Gold Fixings

For a daily value of gold, the London gold fixings from the London Bullion Market Association[1] are used. This is a bi-daily fix of the gold price determined by the five members of The London Gold Market Fixing Ltd; The bank of Nova Scotia, HSBC, Deutsche Bank, Societe Generale Corporate & Investments and Barclays Capital. For over 90 years the gold fixings have been published now, and it is being used as a benchmark for pricing the majority of Gold products.
The fixings are based on an agreement of the 5 members, as to how much they would like to buy or sell gold and in what quantities (how many bars). In case all the members only want to sell or buy, or the number of bars sold/bought does not balance, the price will be moved until an agreement is achieved. The price will be considered fixed if all members agree.

The gold fixings are published at 10:30 AM and 3:00PM (GMT); I will use the morning fixings only, as there are many afternoon fixings that are missing. The gold fixing are conducted in USD, GBP and Euro, I will use USD as standard currency.

[1] http://www.lbma.org.uk/pages/index.cfm

In Figure 2 the gold morning fixings are plotted. One can see that ever since 2003 the gold price has only been rising, apart from a small dip in 2008.

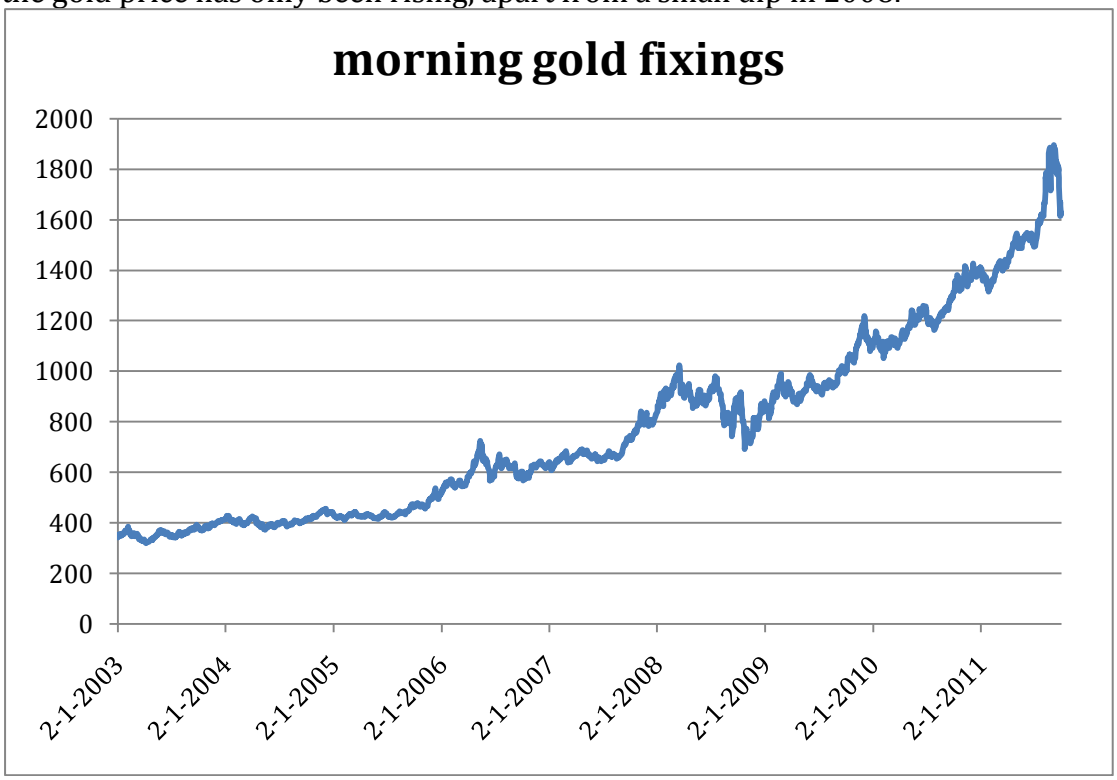


Figure 2: absolute gold fixings

In Figure 3 one finds the log returns of the gold fixings, periods of high and low volatility can be distinguished. Especially in the peak of the credit crunch (September 2008 -2009), we see highly volatile returns.

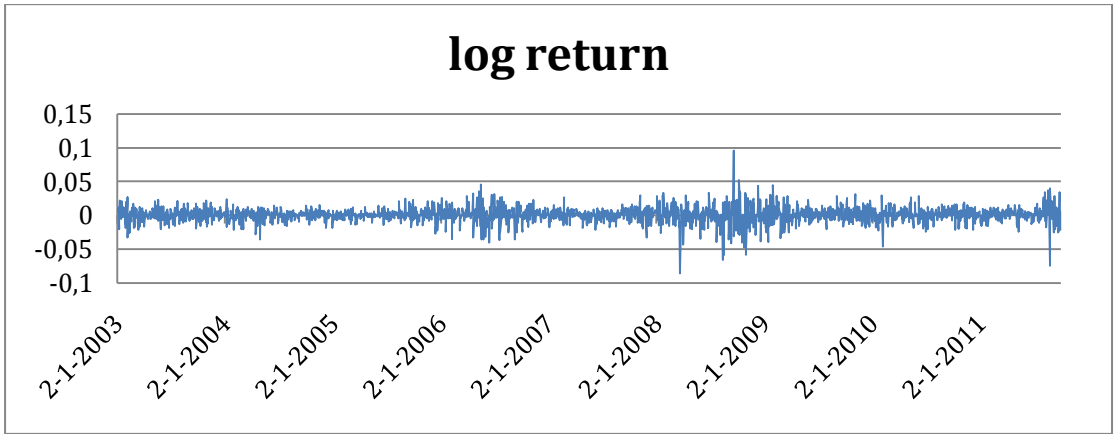


Figure 3: log returns of the gold fixings

Futures

As high frequency data I will use the future contracts on gold. These are contracts are traded on NYMEX. The contract size is 100 ounces of gold. As these are not traded very often (only a few times per day), I will use the bid ask quotes as being an estimation of the true price. I have collected the 10-minute quotes, which equals 144 quotes per day, ranging from January 2007 until September 2011.

More information on these future contracts can be found in appendix B.

In Figure 4 the bid-ask quotes are plotted for the period 2007-2011. As expected we see the same movements as we saw with the gold fixings.

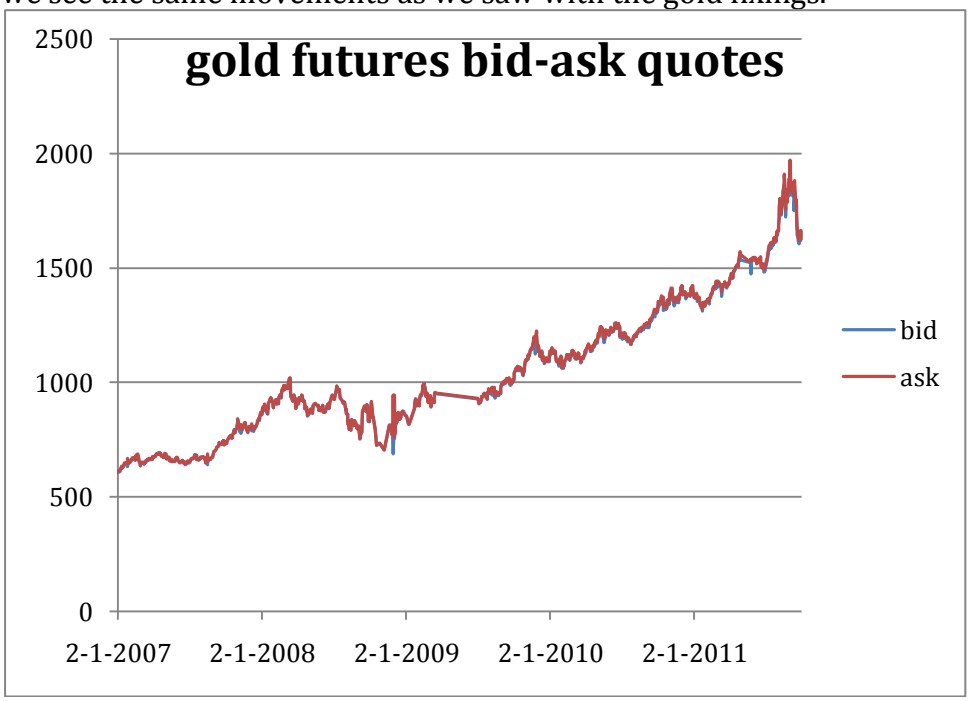


Figure 4: The future bid-ask quotes, ticker code: ZGc1

Aggregation

The news data of Thomson Reuters is delivered to me as separate articles, each with their own sentiment scores. There are a huge amount of articles per day, therefore it is necessary to combine multiple sentiment scores to 1 score. These aggregated scores have much more value than just the sentiment scores of 1 article, i.e. one article will have much less influence on the investor's sentiment than 10. It is important to find a good way to combine, or aggregate, these scores to one score, i.e. the scores have to be aggregated in such a way that they give a good representation of the investor's sentiment caused by the news. The decision of the range of scores that will be aggregated is also very important. To decide on this one has to keep in mind the way peoples memory works. For how long would a human being consider a news article to be relevant?
In Figure 5 a schematic representation of the daily aggregation process is shown. When working on a daily basis, it is assumed that news articles will only influence the first subsequent gold price (return). As there are no prices during the weekend, all the articles that are published between the publication on Friday and the publication on Monday will be aggregated to one score.

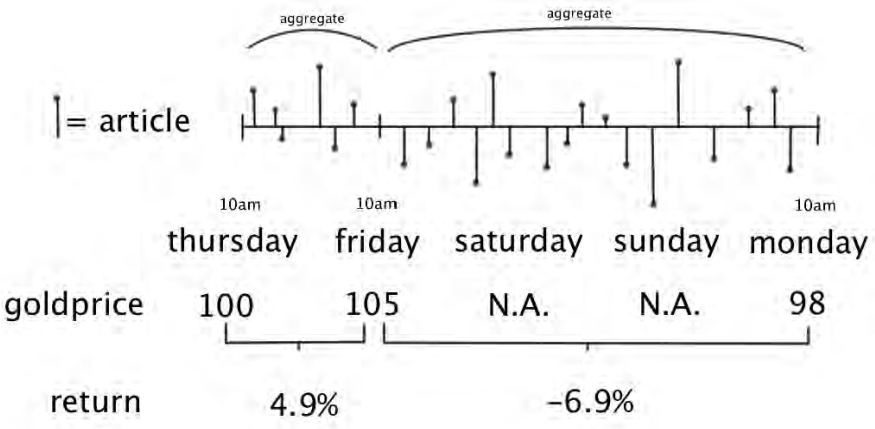


Figure 5: schematic representation of aggregation process

Aggregation methods

There are four methods that will be used to aggregate multiple scores to one score. I will elaborate on them below.

1. Mean

The simplest method is to use the simple (unweighted) mean of the sentiment scores. All the scores that influence the price of a certain period will be averaged to 1 score. The probability scores are averaged without considering the decay of memory. In other words, it is assumed that old news will have the same impact as recent news.

Here the sentiment scores corresponding to each return are averaged. E.g. for a one-day return, the sentiment scores of 1 day are aggregated and for a weekend return the sentiment scores of the whole weekend are aggregated.

2. Bounded

When the negative sentiment and the positive sentiment are almost the same one should consider this noise. The article cannot be clearly specified as positive or negative. Therefore, a boundary for the minimum difference between positive and negative sentiment is introduced. Articles with sentiment scores that exceed this boundary will be omitted from the data. The bound that is used, is the 5% percentile, which means that 5% of the articles (with the smallest differences) will be omitted from the aggregation.

3. Weighted

For the weighted aggregation I introduce two weights for each article: the time weight and the equilibrium weight. The time weights will be used to weight the article and the equilibrium weights will be used to find the weight of the equilibrium. If the articles within the aggregation window are all far in the past, the equilibrium weight will be higher, and the time weight lower. Below I will show the different schemes that I have used to find these weights.

Time weight
As human beings tend to give more value to recent news, some sort of weighting scheme has to be implemented to account for this property. Older news should be given less weight and more recent news should be given more weight. Therefore we will use the increasing weighting schemes show in Figure 6 and Figure 7. The constant c can be adjusted to change the curve/slope of the scheme. The τ is the time difference between the current time and the publication time of an article. The ω are the weights.

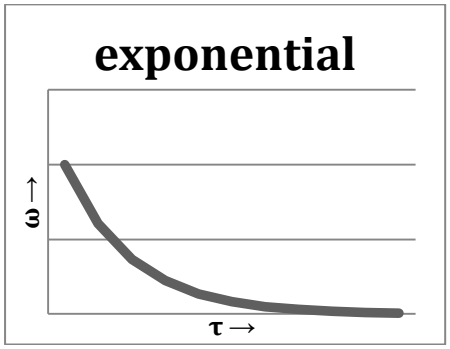


Figure 6:

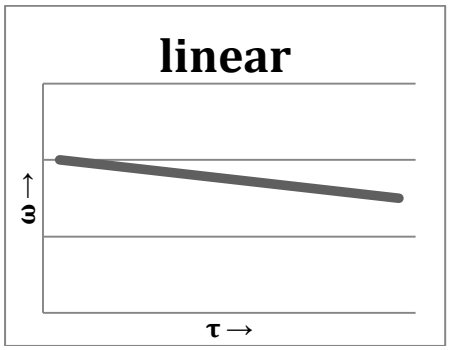


Figure 7: –

Equilibrium weights
Another matter that has to be accounted for is the fact that when there is no news, or only old news, the sentiment score has to move to an equilibrium. In this paper it is assumed that when no news appears, the sentiment scores are one-third, for positive, negative as well as neutral sentiment. When there is no news, it should be assumed that the probabilities for an article to be positive, neutral or negative should be the same; we cannot draw any conclusion as to what the gold price will do according to the news.

When the τ is bigger, more weight should be given to the movement to equilibrium. So the older an article is, more weight is given to the equilibrium as there is more uncertainty as to what the 'real' sentiment at the moment is. Therefore we get the increasing weighting schemes shown in Figure 8 and Figure 9. The constant c can be adjusted to change the curve/slope of the scheme.

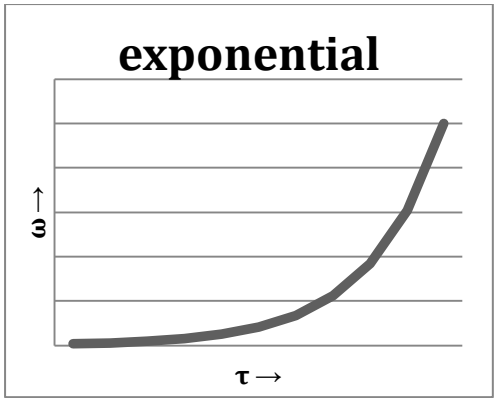


Figure 8: —

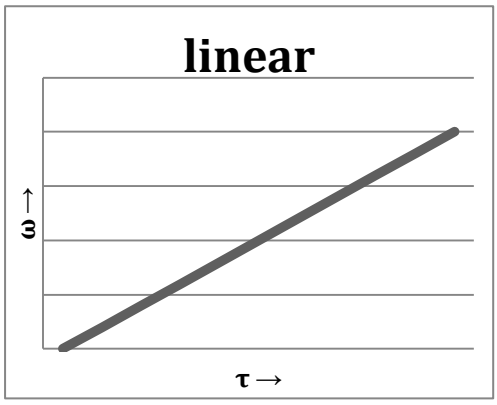


Figure 9: -

4. Weighted+bounded
This is a combination of both methods (bounded and weighted) described above.

Rolling Sentiment

The rolling sentiment gives a sentiment score at each time of the day. Basically, this is a collection of aggregated sentiment scores. Therefore, the value of this rolling sentiment depends on the way the articles are aggregated, and the size of the aggregation window, i.e. how long articles are considered to be relevant.
In this paper, a discrete indication of the general sentiment for gold will be calculated, at each hour of the day. Ideally we would have a continuous rolling sentiment, such that at any moment in time we would be able to produce an index for the sentiment. But considering the scope of this paper, I will only consider discrete, hourly sentiment scores.
I will use a rolling window, with sizes of 1 day and 10 days, and using different weighting schemes to see the differences.

For the different weighting schemes, the rolling sentiment scores are analyzed to find any significant differences. I have used three combinations (Table 2) of the two weight schemes, where the constant c has a value of 5. Also the omission of weights is implemented for comparison.

	Time weights	Equilibrium weights
1	Linear	Exponential

2	Exponential	Linear
3	Exponential	Exponential
4	No weights	No weights

Table 2: weight schemes

In Figure 10 the differences between taking a 1day window and a 10 day window are clearly visible. The 10 day window is a smoothed version of the 1 day window, because it has much more overlapping articles, as the moving window is much larger.

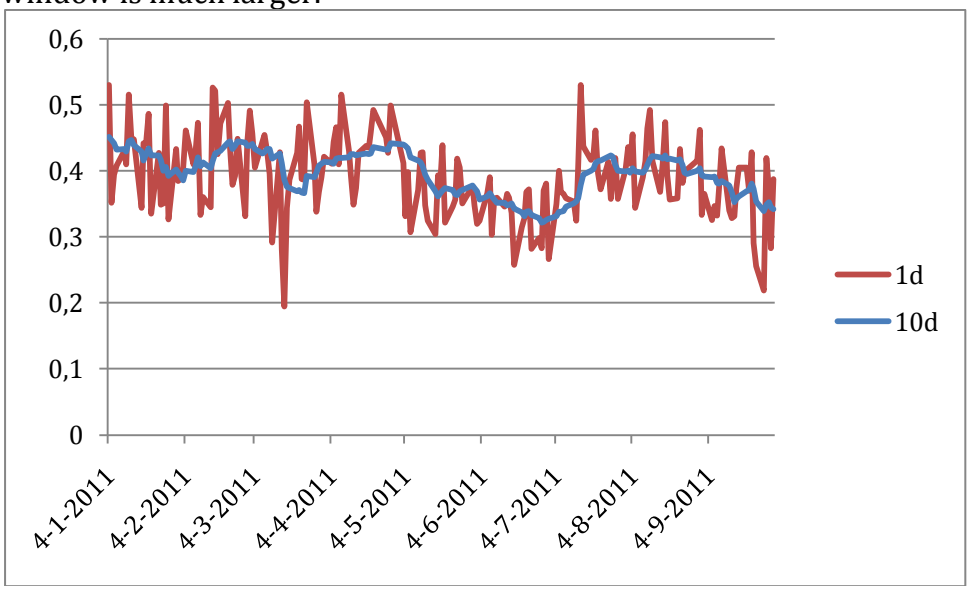


Figure 10: rolling positive sentiment, 10 day window, 4 different weight schemes

The rolling positive sentiment with an aggregation window of 1 day is plotted in Figure 11. This gives a better presentation of the differences between the weighting schemes. As these scores are very volatile we will zoom in on September 2011.

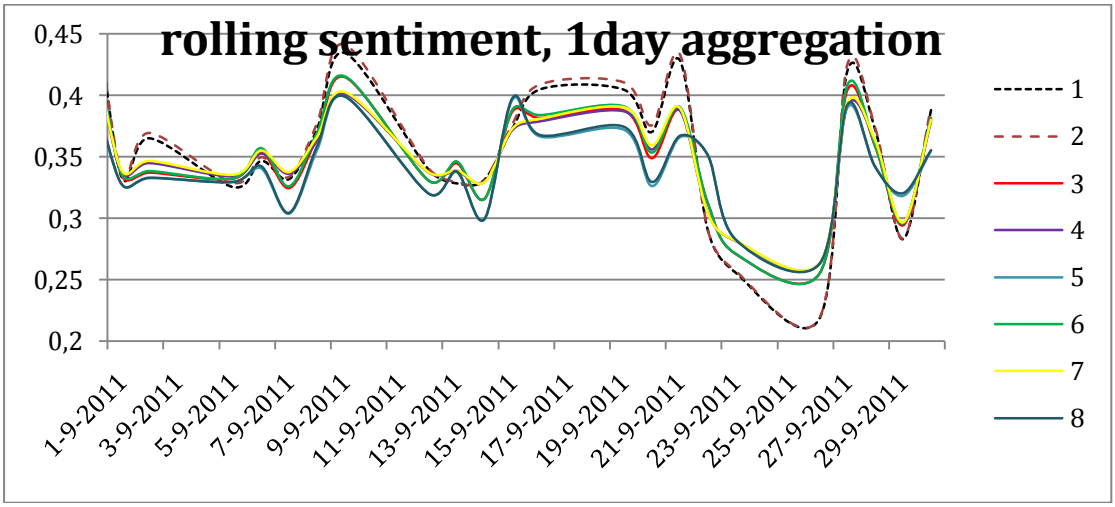


Figure 11: daily rolling sentiment, aggregation window of 1 day, September 2011

In Figure 12 we take a window of 10 days, where we can now clearly see the differences between the 8 aggregation schemes.

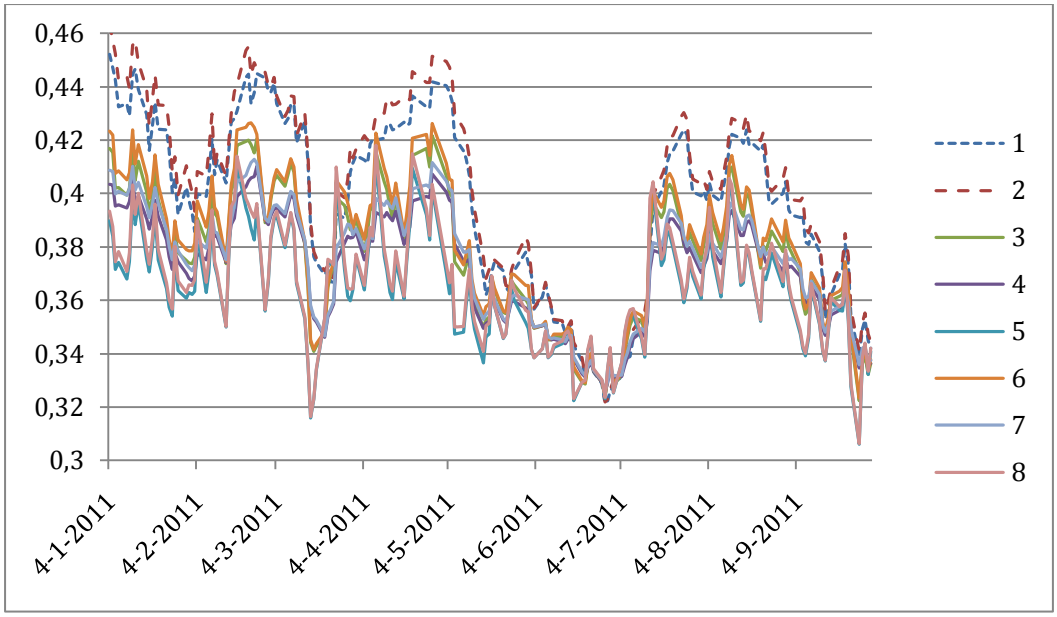


Figure 12: rolling positive sentiment for 2011, 10 day window, 8 different aggregation schemes

When using no weights we find the highest rolling sentiment and the twice exponential scheme gives us the lowest rolling sentiment scores.
In this small sample, one can see that when using a different weighting scheme combination, the rolling sentiments approximately move in the same way.

Correlations

One can measure the news-gold relationship by using the correlation coefficients of the daily returns with the aggregated sentiment scores for the 4 methods. I will use three different combinations of weighting schemes shown in Table 3, as well as a constant parameter c=5. I will give these combinations number 1, 2 and 3 when throughout this paper.

	Time weights	Equilibrium weights
1	Linear	Exponential
2	Exponential	Linear
3	Exponential	Exponential

Table 3: 3 different weight schemes

In Table 4 the correlations of the daily aggregated sentiment scores are shown with a rolling window of 1 day.

Daily correlations

	Sent_pos	Sent_neut	Sent_neg
mean	0.4422	0.0263	-0.4657
Weighted 1	**0.4944**	**-0.0022**	**-0.5285**
Weighted 2	0.4842	0.0112	-0.5148
Weighted 3	0.4558	-0.0246	-0.4964
Bounded	0.4425	0.0262	-0.4665
Weighted+bounded 1	0.4933	-0.0021	-0.5292
Weighted+bounded 2	0.4839	0.0116	-0.5156
Weighted+bounded 3	0.4534	-0.0237	-0.4966

Table 4: correlations of aggregated sentiment scores with gold fixing returns, 1 day aggregation window

It is very pleasant to see that the one day correlations are really high. This means that the news sentiment scores and the gold fixing returns move in similar ways. The aggregation window of 1 day seems to be chosen well.
In Table 5, the correlations of the daily aggregated sentiment scores are shown with a rolling window of 10 days.

	Sent_pos	Sent_neut	Sent_neg
mean	0.0757	0.0359	-0.1041
Weighted 1	0.1305	0.0246	-0.1641
Weighted 2	0.0912	0.0290	-0.1205
Weighted 3	**0.2344**	**0.0142**	**-0.2797**
Bounded	0.0756	0.0354	-0.1054
Weighted+bounded 1	0.1306	0.0233	-0.1650
Weighted+bounded 2	0.0913	0.0279	-0.1214
Weighted+bounded 3	0.2337	0.0128	-0.2801

Table 5: correlations of aggregated sentiment scores with gold fixing returns, 10 day aggregation window

One can see that when increasing the rolling sentiment window, the correlations are significantly lower. This is intuitively, as 10 day old news will most probably have no influence on the current gold price. In 10 days there have been published so many new articles, the oldest ones become insignificant and should thus not be included in the aggregation.

High frequency correlations

	Sent_pos	Sent_neut	Sent_neg
mean	0.0647	0.0096	-0.0736
Weighted 1	0.0468	0.0243	-0.0681
Weighted 2	0.0507	0.0208	-0.0694
Weighted 3	0.0303	0.0276	-0.0559
Bounded	**0.0660**	**0.0081**	**-0.0740**
Weighted+bounded 1	0.0476	0.0238	-0.0687
Weighted+bounded 2	0.0515	0.0205	-0.0701
Weighted+bounded 3	0.0309	0.0274	-0.0566

Table 6: correlations of aggregated sentiment scores with gold future returns, 1 day aggregation window

On a high frequency basis, the correlations are considerably smaller. There is too much noise in high frequency data to find any correlation.

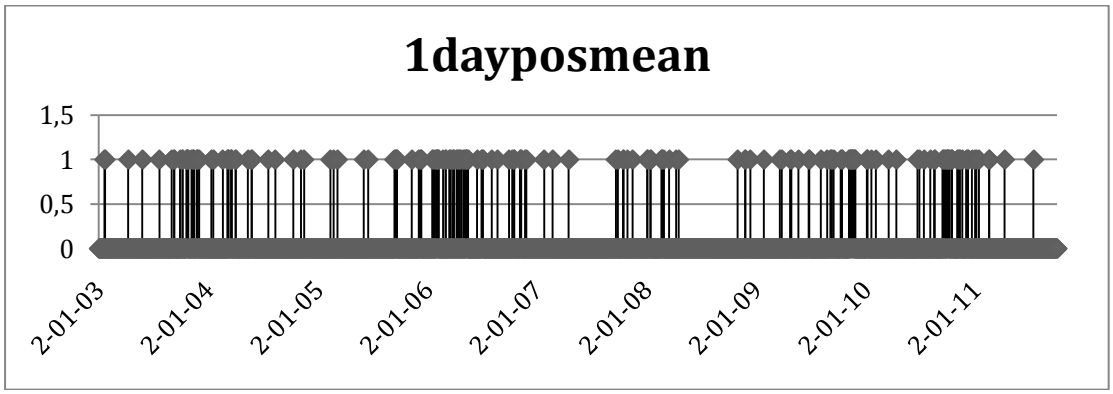
Event studies

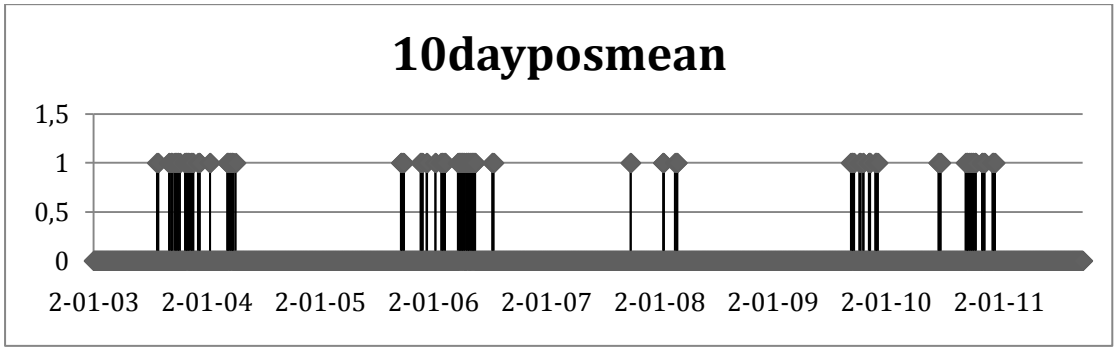
Event studies are widely used to find patterns in the returns that occur at certain 'events.' In this paper the events are extreme news sentiments, both positive and negative. A certain percentage of the most negative and positive aggregated sentiment scores are considered events. And the corresponding dates are called event dates. The cumulative returns are analyzed around these event dates. As a result, it gives us some insight in the way these events influence the returns of commodities, in this case gold. Let us first take a look at the way these event dates are distributed over our timeframe (2003-2011). We will compare

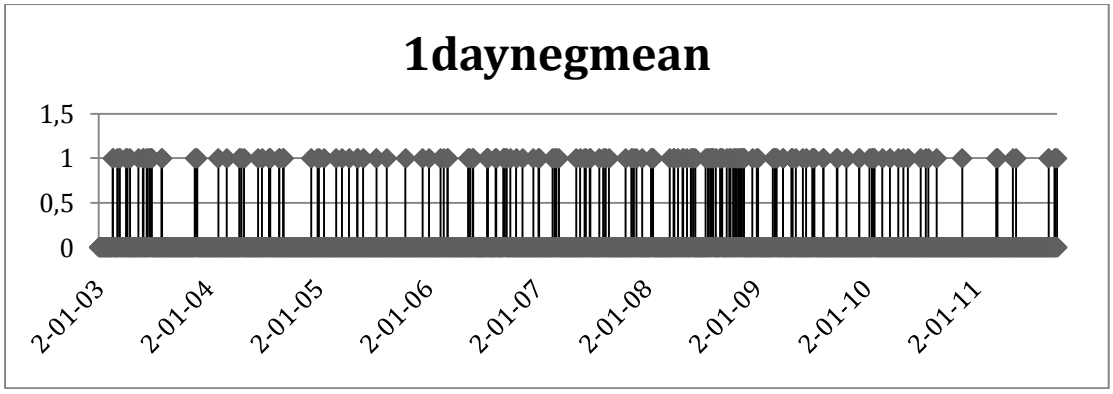
different aggregation windows, to see the influence that various window sizes have on the dates that events occur.

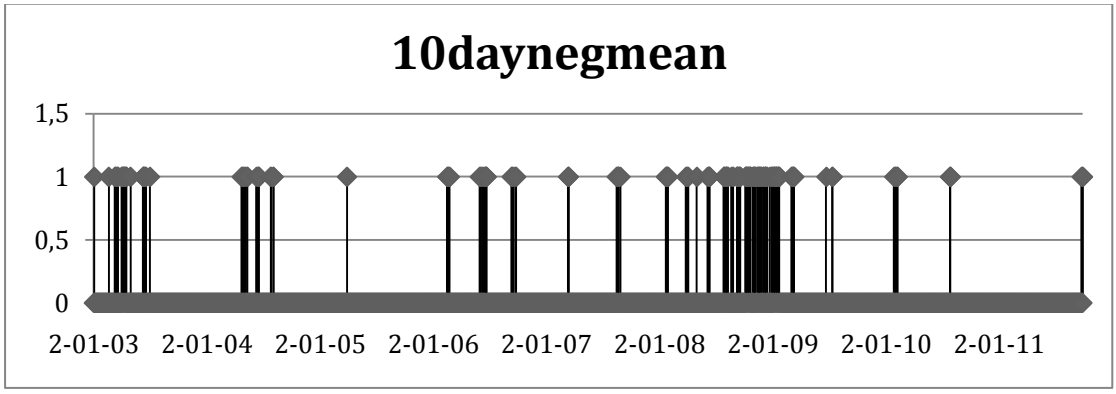
Event dates

To start the event studies one should start by looking at the times that events take place. For the dataset that I use (2003- September 2011) the event dates are plotted. Here the scores are aggregated once a day at 10:30AM, when the gold fixings are published. An aggregation window of one day and an event bound of 10% (so the 10% most negative and 10% most positive aggregated sentiment scores) is used here.

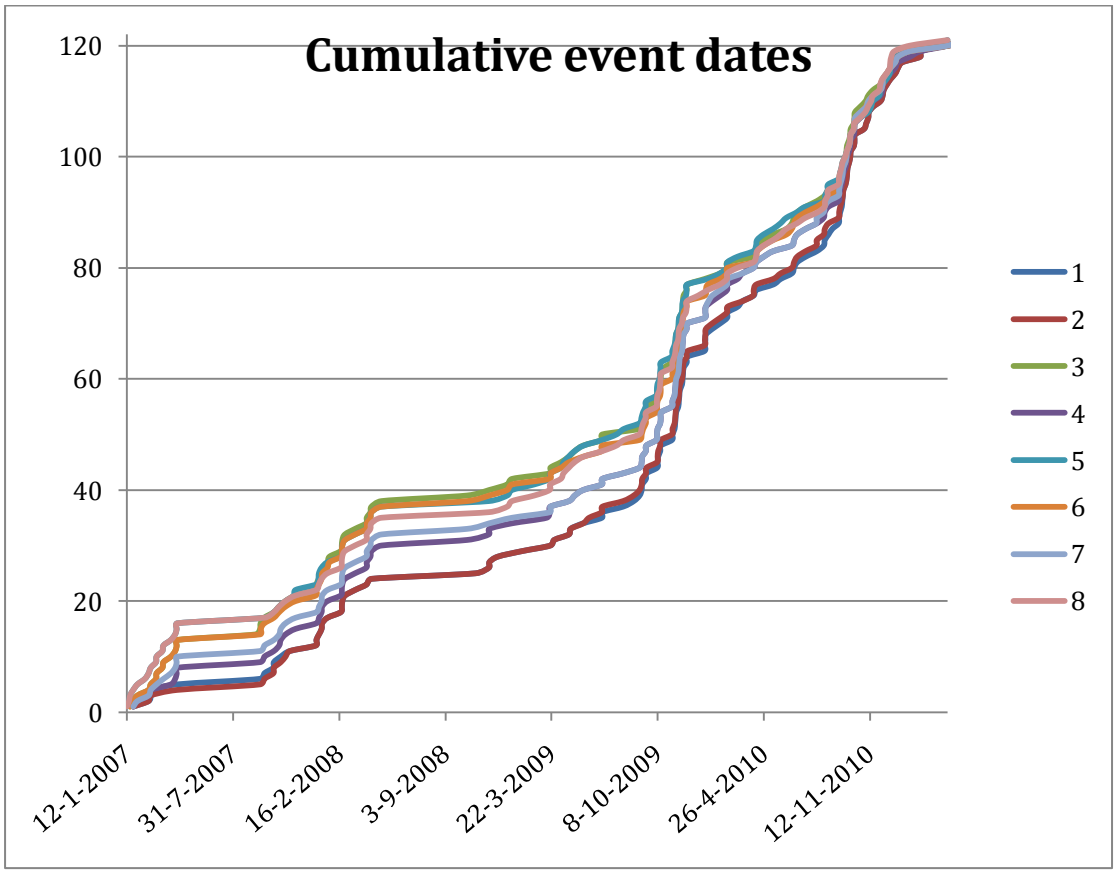








We can see that when using a 10 day window to aggregate the articles, the events are much more clustered.



Here the cumulative events are plotted, and we can clearly see that they move in the same way. This indicates that using different combinations of weighting schemes doesn't make much of a different in the event studies. The weighting schemes are used to create a more realistic score, but in fact don't really differ much from each other.

Event studies on daily data

For each aggregation method (see above) the event dates are found, and with a window of 41 days [-20, +20] the cumulative returns are calculated. The event occurs at time 0. Missing returns are linearly interpolated with the available prices; e.g. when a return of Thursday is missing, this will be calculated as ½ times the return of Wednesday till Friday. Each figure contains 24 lines, corresponding to the combinations of 4 methods used for aggregation, the 3 different weighting schemes that were discussed in the previous chapter and the three types of events (positive, neutral and negative). The neutral events are added for comparison.
A legend of the 24 lines is given in Table 7.

	Aggregation method	type	color
1	Mean	Positive	
2	Mean	Neutral	
3	Mean	Negative	
4	Bounded	Positive	
5	Bounded	Neutral	
6	Bounded	Negative	
7	Weighted1	Positive	
8	Weighted1	Neutral	
9	Weighted1	Negative	
10	Weight+bound1	Positive	
11	Weight+bound1	Neutral	
12	Weight+bound1	Negative	
13	Weighted2	Positive	
14	Weighted2	Neutral	
15	Weighted2	Negative	
16	Weight+bound2	Positive	
17	Weight+bound2	Neutral	
18	Weight+bound2	Negative	
19	Weighted3	Positive	
20	Weighted3	Neutral	
21	Weighted3	Negative	
22	Weight+bound3	Positive	
23	Weight+bound3	Neutral	
24	Weight+bound3	Negative	

Table 7: legend of Figure 13

Where the weighted schemes are number according to Table 3.

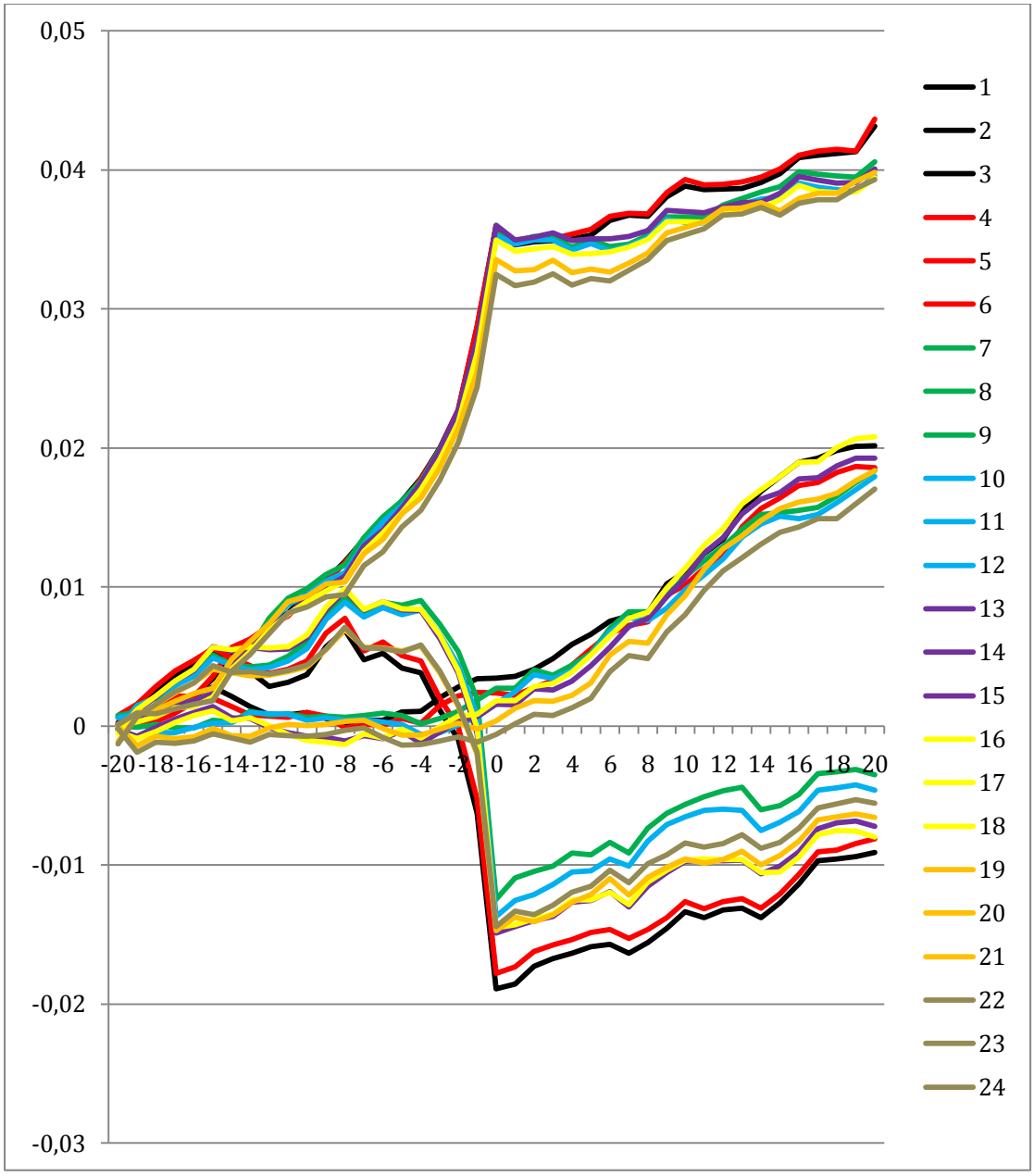


Figure 13: Event studies on daily aggregated data, cumulative returns on the window [-20,20]

It is clearly visible that just before the event the cumulative returns take a huge leap, this is negative for negative events and positive for positive events. After the events we see the cumulative returns move back to the equilibrium. This would indicate that negative news is a reaction to large negative movements in the gold price, and for positive news vice versa. There are not very large differences between the different aggregation methods.
By adding confidence bounds, one can get a measure of the uncertainty of these returns around the events.
The confidence bounds are given by

For a 95% normal confidence interval.
In Figure 14 the mean aggregated events (no weights), the confidence bounds for the cumulative returns are added.

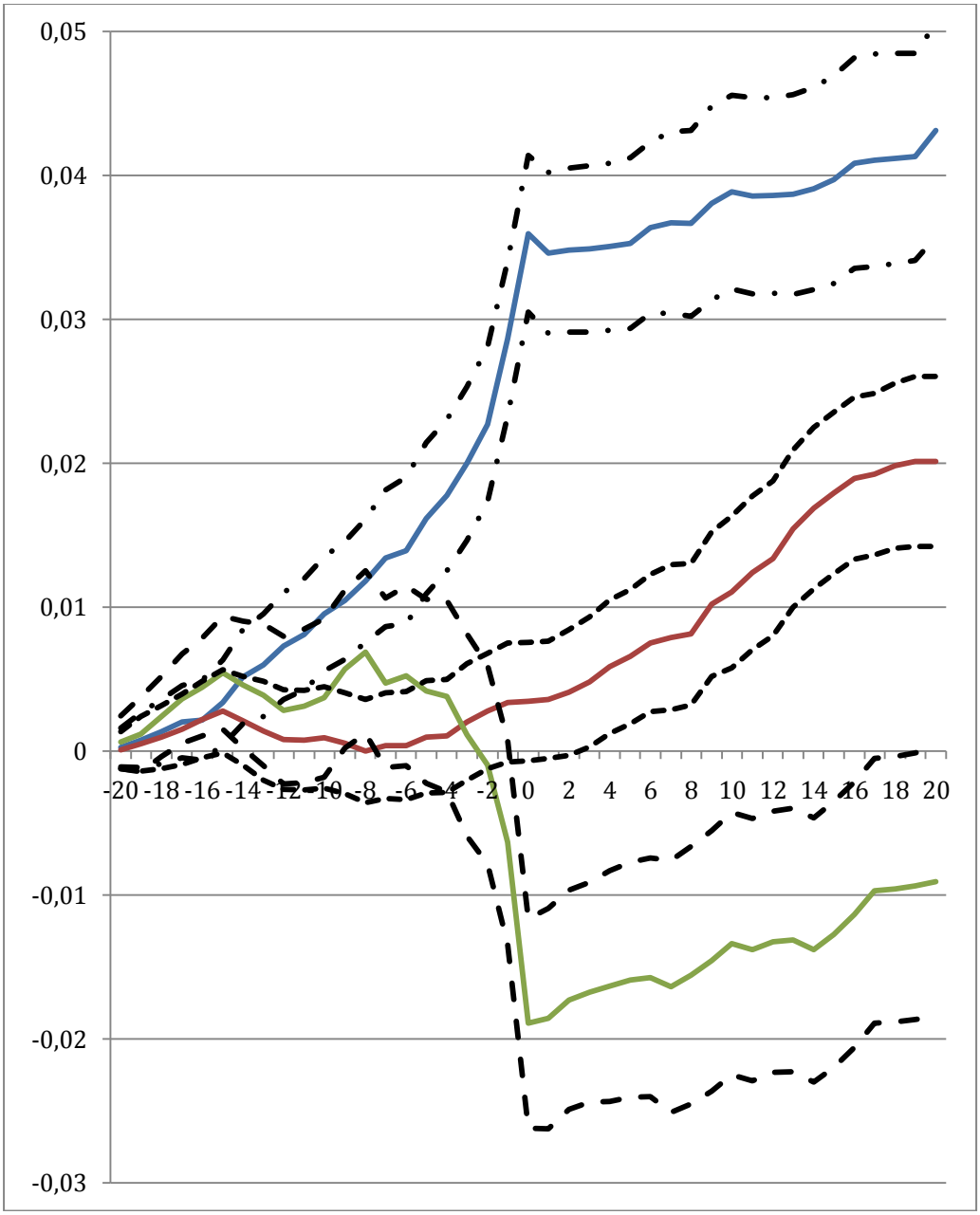


Figure 14: cumulative returns of the mean aggregated event dates, including confidence bounds

Event studies on intraday data

For these event studies the futures data I collected from Reuters is used. As there are not many transactions in these products, the bid and ask quotes are used to give an estimation of the future prices at each point in time. In Figure 15 and 16 the bid and ask quotes are plotted for positive and negative events, respectively.

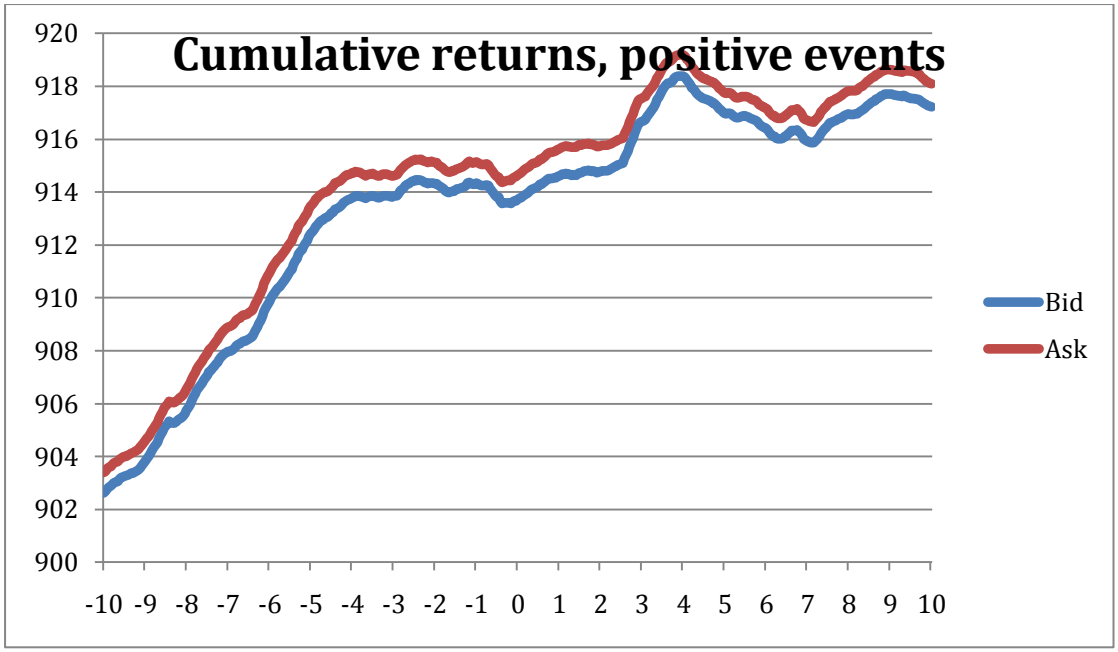


Figure 15: cumulative returns around positive events, events found on intraday base

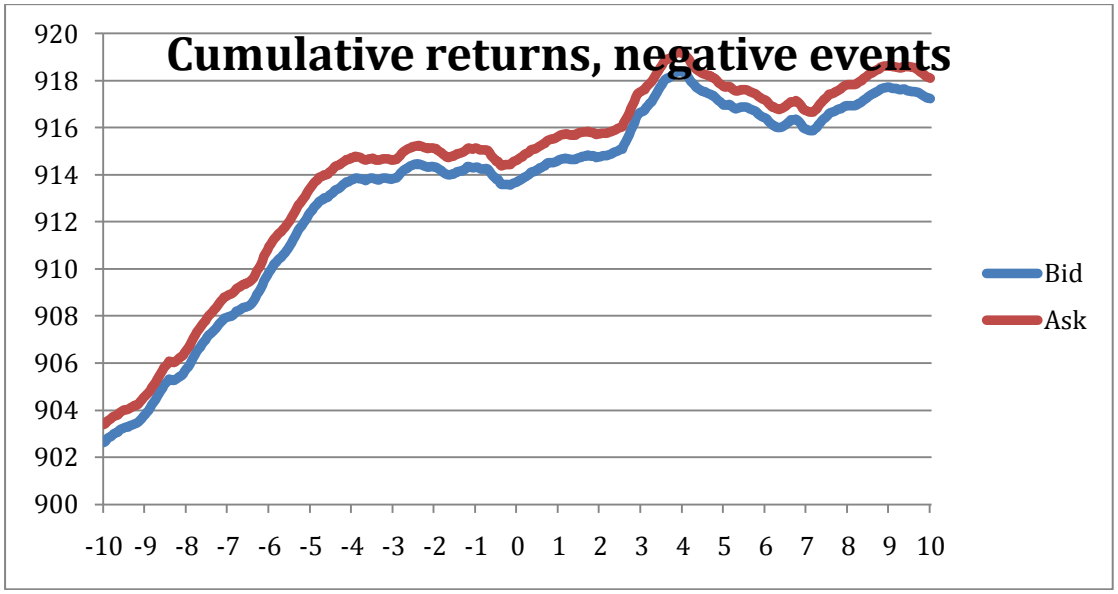


Figure 16: cumulative returns around negative events, events found on intraday base

In the intraday event studies, no clear movements can be seen around the event dates. The reason for this is that this data probably contains too much noise to get clear results. Therefore, I will perform an investment strategy using these bid-ask quotes, to see whether a profit can be made by trading at high sentiment points.

Investment Strategy

The ultimate goal is to find a strategy that can generate a profit from this rolling sentiment score. In order to find out if this is possible, an investment strategy, based solely on the news sentiment scores, is implemented. I will use some sort of event studies to find explicitly high news sentiment scores. The rolling sentiment explained in one of the previous chapters will be used for this.

Strategy

The strategy that I will implement will be a strategy that trades at moments of extreme sentiment, or 'events'. To determine when an event occurs we need to set the strategy bounds. A bound for positive sentiment and for negative sentiment is set. When these bounds are found, the strategy can be executed. Each time the rolling sentiment crosses the strategy bounds, gold will be bought for the current ask quote. Then for the subsequent 10 days, I will see what the best time to sell the future (for the current bid quote) is.

Setup

First one has to separate the data into two datasets; one for setting the strategy parameters and one for the execution of the strategy.
I will use the data from 2007-2010 for setting the strategy parameters, i.e. the bounds for positive and negative events. I will use the 10% quantiles as strategy bound, both for positive and negative sentiment scores. The scores are aggregated with a windowsize of 1 day, and with 8 different aggregation methods shown in Table 8.

For the testing of the strategy I will use the most recent data that I have, that is the data from January 2011 until September 2011.
I have also calculated a benchmark, which is the annualized profit of investing in gold at the start of 2011 until the end of September 2011. The ask quote at the start of 2011 is 1376.80 USD and the last ask quote of September 2011 is 1634.40 USD. So we set the benchmark at —— — . Which is an estimated profit of 22.87% over 2011.

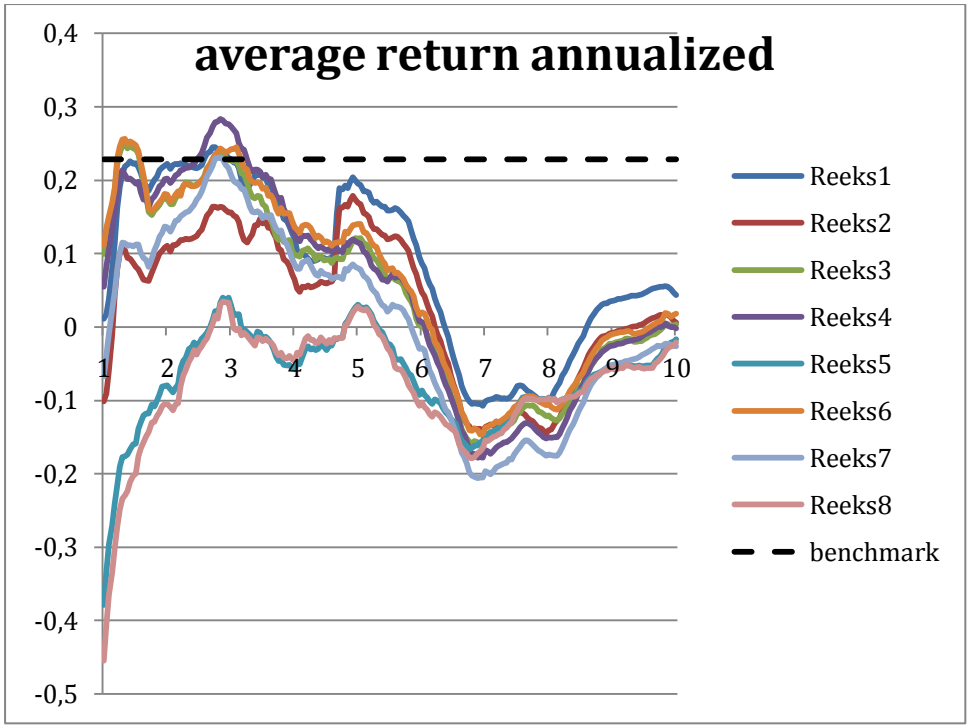
Results

I have compared 8 different aggregation methods. Some of them produce a profit and some don't. In Table 8 the aggregation schemes are shown.

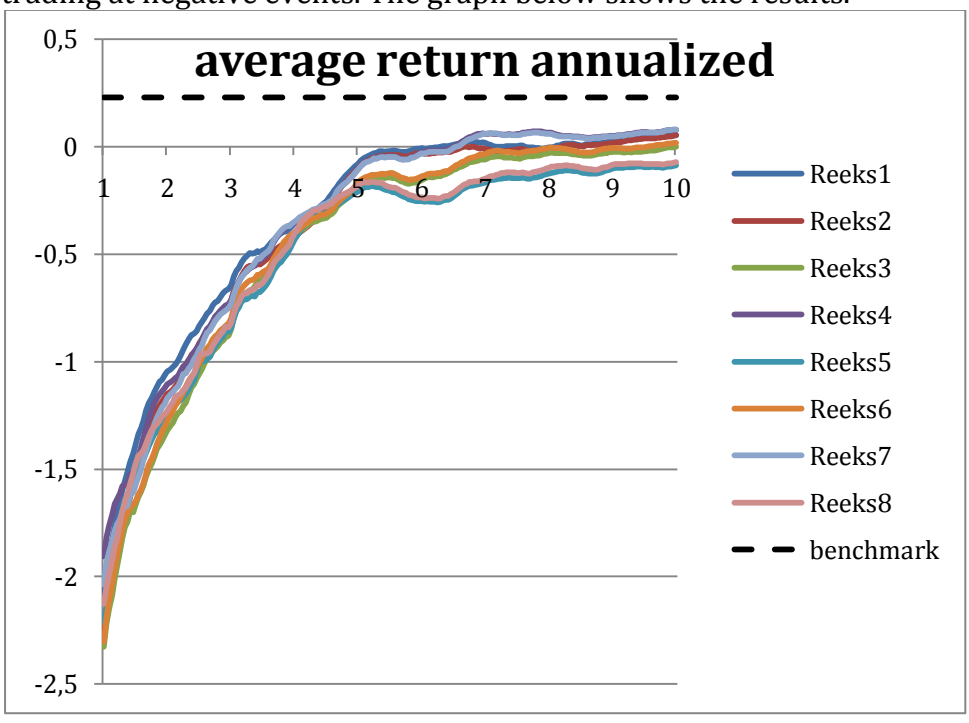
	method	Time weights	Equilibrium weights
1	Mean	N.A.	N.A.
2	Bounded	N.A.	N.A.
3	Weighted	Linear	Exponential
4	Weighted	Exponential	Linear
5	Weighted	Exponential	Exponential
6	Weight+bound	Linear	Exponential
7	Weight+bound	Exponential	Linear
8	Weight+bound	Exponential	Exponential

Table 8: aggregation methods for the strategy

Here the strategy is shown for the 8 different aggregation schemes. This strategy trades gold at an event. And one can see that the maximum is for series 4, which is exponentially linearly weighted aggregation, for a short term investment period of 3 days. The benchmark shows the annual profit when buying gold futures at the start of 2011 and selling them at the end of 2011 (without considering the rollover costs). When investing 3 days by using the strategy bounds found with exponentially linearly weighted aggregation, we find an average annualized return of 27.28%. This is a difference of +4.41 percentage point with the benchmark.



To compare the result I have implemented the same strategy as before, but now trading at negative events. The graph below shows the results.



It is clearly visible that investing at negative events is not a good idea. None of the aggregation schemes give a better result than the benchmark. This shows that the events do have an effect on the price movements of gold futures.

Conclusion

The Thomson Reuters News Analytics Engine is a highly sophisticated news reading engine. Therefore, investigating this engine has been a thrill and I have found some interesting results in the relationship between the sentiment scores and the gold price. I was left very free as how to interpret the data, and therefore it was hard at times to set the scope of this paper. As there are so many different ways to interpret the data. In the end I used 8 different ways to aggregate the sentiment scores, which include three different weighting schemes.
Both daily and intraday gold data is used in this paper.
First I looked at the simple correlations between gold returns and the aggregated sentiment scores. For the daily data these correlations were very high. But unfortunately, the future quotes contain too much noise, so produced a very low correlation with the aggregated sentiment scores.
This also caused the event studies to fail to show proper results on an intraday basis. On a daily basis, the event studies did show heavy movement in the cumulative returns around event. This movement starts around four days before the event occurs and lasts until the moment of the event. This indicates that prices aren't driven by the news, but the news is driven by the prices. When looking at actual gold news, a lot of articles are about the price of gold. So consequently these event study results are found.
But the main question is, is there a way to make a profit using only the data from the news sentiment engine? To answer this question I have implemented a strategy that, considering the analysis of the data, seemed adequate. The strategy trades gold when the sentiment score is exceptionally high. In fact it trades gold when the sentiment scores are in the upper 10% percentile. The results of the strategy were surprisingly good. Even when the event studies didn't show any positive movements after events, the strategy for 2011 produced 4.41 percentage points better than the benchmark. Where the benchmark is the profit made by investing a full year (2011) in gold.

Future research

The news sentiment engine has so many implementations. There are a large number of assets and commodities that are available in this engine.

In this research I have not taken the volume of the news into consideration. One could, for example, model the volatility by the volume of the news articles. As more news would probably cause a higher volatility. Many other implementations are possible considering volatility model.

One of the problems I encountered was that the high frequency data on gold was very hard to find. Therefore I used the bid-ask quotes that we found, but as I found out, this data is very noisy. When having access to high frequency gold data, maybe from companies that trade in gold, a lot of doors would be opening. As prices react to news on a very high frequency basis.

These are just a few things that are possible with this news engine, as this field of research is extremely broad.

Appendix A

Number of articles per month:

	2003	2004	2005	2006	2007	2008	2009	2010	2011	
jan	1894	2212	1576	1968	3066	3080	2385	2830	1596	20607
feb	1853	2413	1674	3296	2899	3085	2838	3084	1665	22807
mar	1609	2805	1710	3564	2915	3014	3002	3096	2223	23938
apr	1582	2521	1494	3664	2605	2709	2728	3050	2098	22451
may	2191	2181	1525	4360	3090	2655	2460	3380	2473	24315
jun	2346	2731	1780	3792	2828	2588	2499	2719	2224	23507
jul	2729	2544	1590	3107	3234	2869	2921	2430	2810	24234
aug	2414	1597	1929	2777	2609	2982	2345	2459	3738	22850
sep	2643	1723	1884	2718	2884	3395	3499	3390	2971	25107
oct	2159	1628	1769	2796	3307	3265	3896	3560	0	22380
nov	2451	1657	1904	3124	2995	2350	4380	3580	0	22441
dec	2356	1562	1754	2586	2132	2520	3060	2805	0	18775
	26227	25574	20589	37752	34564	34512	36013	36383	21798	273412

positive sentiment: 122894
neutral sentiment: 36983
negative sentiment: 113535

Appendix B

U.S. MARKET
PRECIOUS METALS COMPLEX

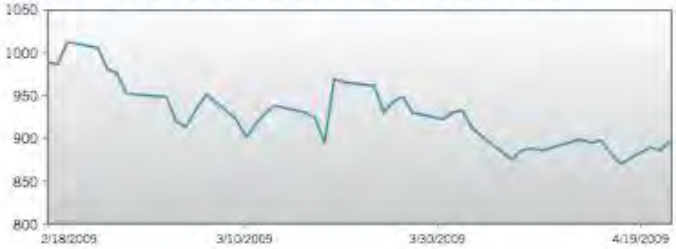
100 Ounce Gold Futures

NYSE Liffe U.S., NYSE Euronext's U.S. futures exchange launched on September 8, 2008, provides a fully electronic, liquid market for physically deliverable 100 ounce gold futures, 5,000 ounce silver futures, options on gold and silver futures, and mini-sized 33.2 ounce gold and 1,000 ounce silver futures. Our metals complex is the foundation of our world class, multi-asset futures exchange providing global customers significant value and enhanced trading opportunities. NYSE Liffe U.S. leverages a broad range of risk management products, global distribution network and leading edge LIFFE CONNECT 10® technology.

100 Ounce Gold Futures and Options on NYSE Liffe U.S.

100 Ounce Gold Futures Contract	
Trading Unit	100 fine troy ounces
Price Quote	U.S. dollars and cents per ounce
Tick Size	$0.10 per ounce; $10 per contract
Contract Months	Current month (for delivery purposes) and the next 2 months. Any February, April, August and October within 23 months of the current month. Any June and December within 60 months of the current month.
Delivery	New York Exchange approved vaults. 100 fine troy ounces (±5%), no less than 99.5% fineness, consisting of one 100 ounce bar or three 1 kilogram bars.
Last trading session without delivery risk	Long positions must be offset prior to the close of the trading session ending on the second business day prior to the first business day of the contract month.
Last Trading Day	Third to last business day of contract month at 1:30pm U.S. ET
First Delivery Day	The first business day of the contract month.
Last Delivery Day	The last business day of the contract month.
First Seller's Notice Day	First day short can notify Exchange of intent to deliver is the second business day prior to the first delivery day.
Last Seller's Notice Day	Last day short can notify Exchange of intent to deliver is the second business day prior to the last delivery day.
First Buyer's Notice Day	First day Exchange can notify the long has been assigned the obligation to make payment for and take delivery of an electronic vault receipt is the business day immediately prior to the first delivery day.
Last Buyer's Notice Day	Last day Exchange can notify the long has been assigned the obligation to make payment for and take delivery of an electronic vault receipt is the business day immediately prior to the last delivery day.
Regular Trading Hours	7:16pm to 5:00pm the following day, Sunday through Friday, U.S. ET
Regular Daily Settlement	1:30pm U.S. ET

NYSE Liffe U.S. 100 Ounce Gold Futures Settlement Price



NYSE Liffe U.S.

References

http://www.lbma.org.uk/pages/index.cfm?page_id=1

http://www.goldfixing.com/

http://www.nyse.com/