Association for Information Systems
AIS Electronic Library (AISeL)

ECIS 2010 Proceedings

European Conference on Information Systems (ECIS)

2010

Enhancing Automated Trading Engines To Cope With News-Related Liquidity Shocks

Sven S. Groth
E-Finance Lab Frankfurt, sgroth@wiwi.uni-frankfurt.de

Follow this and additional works at: http://aisel.aisnet.org/ecis2010

Recommended Citation
Groth, Sven S., "Enhancing Automated Trading Engines To Cope With News-Related Liquidity Shocks" (2010). *ECIS 2010 Proceedings*. 111.
http://aisel.aisnet.org/ecis2010/111

This material is brought to you by the European Conference on Information Systems (ECIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2010 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

ENHANCING AUTOMATED TRADING ENGINES TO COPE WITH NEWS-RELATED LIQUIDITY SHOCKS

Journal:	*18th European Conference on Information Systems*
Manuscript ID:	ECIS2010-0367.R1
Submission Type:	Research Paper
Keyword:	Text mining, Decision support systems (DSS), E-finance, Forecasts and implications

ENHANCING AUTOMATED TRADING ENGINES TO COPE WITH NEWS-RELATED LIQUIDITY SHOCKS

Groth, Sven S., E-Finance Lab, Campus Westend, Grüneburgplatz 1, House of Finance, 60323 Frankfurt am Main, Germany, sgroth@wiwi.uni-frankfurt.de

Abstract

Liquidity constitutes one of the main determinants of implicit transaction costs. Deriving optimal execution strategies that minimize transaction costs, automated trading engines need to forecast future liquidity levels. By means of an empirical study we provide evidence that the publication of regulatory corporate disclosures is followed by abnormal liquidity levels. As we do not find abnormal liquidity levels prior to the publication, we assume the content to be largely unanticipated. Forecasting models purely based on quantitative input data may therefore not be able to pick up on the liquidity trends in a timely manner. Against this background, we propose two trading signals that allow automated trading engines to appropriately react to news-related liquidity shocks: First, a simple binary "news" or "no news" signal. Second, a signal that indicates whether or not the publication of a regulatory corporate disclosure will be followed by a negative liquidity shock. Utilizing text mining techniques, the content of the corporate disclosures is analyzed to extract the trading signal. The trading signals are evaluated within a simulation-based use case and turn out to be valuable. We strongly advise developers of automated trading engines to integrate unstructured qualitative data into their models, i.e. the proposed trading signals.

Keywords: Automated trading, liquidity, forecasting, text mining, e-finance, and simulation.

1 INTRODUCTION

The evolution of electronic order books eased the way for the automation of trading processes: Today, the group of computer-based *automated*[1] *traders* generates already about one-half of trading activity on major European markets such as Deutsche Börse's Xetra and the percentage share continues to grow (Deutsche Börse AG 2008). Computer-based *automated traders* "emulate a broker's core competence of slicing a big order into a multiplicity of smaller orders and of timing these to minimize market impact" (Gomber and Gsell 2006). We adhere to the provided definition with regard to the fact that the primary task of *automated traders* is to execute orders, received from the buy side, at best available conditions. In this context, *automated trading* is not about itself finding investment opportunities by for instance exploiting some kind of market imperfections.

Deriving an optimal execution strategy for a pre-defined execution time period, *automated traders* constantly need to trade off the evolving transaction cost components. Associated costs to implementing investment decisions can basically be divided into two broad categories: First, explicit costs such as commissions, fees or taxes. Second, implicit costs such as market impact, timing costs or opportunity costs (Bikker, Spierdijk, Hoevenaars and van der Sluis 2006). Especially for large trades, implicit trading costs are usually much larger than explicit trading costs. Liquidity constitutes one of the main determinants of incurred market impacts, i.e. of implicit trading costs. Therefore, to derive an optimal execution strategy and to minimize associated transaction costs, the ability to forecast future liquidity levels is vital.

Bikker et al. (2006), however, conclude that "forecasting market impact costs appears notoriously difficult and traditional methods fail". Moreover, Domowitz and Yegerman (2005) find that the execution quality of automated traders is inferior to the executions handled by a broker. One possible reason for this may be found with the fact that currently employed models are solely based on purely quantitative input data. Existing (academic) models and strategies largely neglect one of the most important sources of information: unstructured qualitative data, i.e. news (Coggins, Lim and Lo 2006). If, for example, a listed company issues an unanticipated regulatory ad hoc disclosure, *automated traders* cannot sufficiently fast react to it simply because these cannot read (and interpret) the content. As unanticipated news – per definition – are very unlikely to be reflected in time series data prior to publication date, *automated traders* can only respond to other market participants' reactions.

Against this background, we propose two trading signals that provide an indication of expected liquidity levels subsequent to the publication of regulatory corporate disclosures. Extracting the trading signals, we utilize sophisticated text mining techniques. The signals are intended to be used inside existing models. The rest of the paper is structured as follows: First, we present related finance and information systems research. Second, we provide the reader with valuable insights into liquidity dynamics subsequent to the publication of regulatory corporate disclosures. Third, building upon these insights we propose a text mining approach that predicts the liquidity impact of ad hoc news. The classification quality, i.e. the extracted trading signal, is tested by both *classic* model evaluation metrics and domain-specific *simulation-based* model evaluation. Finally, we conclude.

2 RELATED WORK

Besides the literature already mentioned in the previous section, this paper is strongly related to the following streams of finance research and information systems research.

[1] The terms *automated trading* and *algorithmic trading* are used interchangeably in this study.

The effect of information arrival on company-specific liquidity levels has already been analyzed in previous studies (e.g. Lee, Mucklow and Ready 1993). Most of the contributions, however, merely concentrate on one event type such as earnings announcements (Krinsky and Lee 1996) or dividend announcements (Graham, Koski and Loewenstein 2006). Moreover, authors use liquidity measures that do not allow a complete analysis of execution costs (Ronaldo 2003). Finally, other studies are based on very few events and / or short time periods (Brooks, Patel and Su 2003; Gomber, Schweickert and Theissen 2004). We address all of above shortcomings in this paper: We analyze a comparatively large dataset that consists of several regulatory news types. Being provided with high-frequency order book data, we apply a liquidity measure that is especially suitable for the estimation of implicit transaction costs (i.e. liquidity). Finally, we focus on short-term intraday liquidity effects.

The prediction of company-specific liquidity levels in general and the prediction of the liquidity impact of regulatory corporate disclosures in particular have – to our knowledge – not yet been addressed by utilizing text mining techniques. There exists, however, a stream of research that is dedicated to the application of data / text mining techniques in financial markets. Mittermayer and Knolmayer (2006) provide a good survey on existing text mining systems: Since Wuthrich, Cho, Leung, Permunetilleke, Sankaran and Zhang (1998) proposed one of the first financial text mining applications, systems were refined by focusing on aspects such as intraday data, (new) data mining techniques, other forecasting objects (stock prices, exchange rates, volatility), news types (ad hoc news) or novel evaluation methods (Groth and Muntermann 2009). Our approach builds upon this literature and uses intraday high-frequency data on the new forecasting object *liquidity*. Moreover, we concentrate on a certain news type and develop a novel domain-specific evaluation metric.

3 INTRADAY LIQUIDITY DYNAMICS INDUCED BY NEWS

3.1 Dataset

The news dataset at hand consists of ad hoc disclosures published by *Deutsche Gesellschaft für Ad-hoc-Publizität* (DGAP) on behalf of the companies admitted to trading on an organized market in Germany. To fulfil legal requirements, these companies have to publish immediately any insider information or other information being highly relevant to investors. We concentrate on this news type because the disclosures are expected to primarily contain new information and event studies have shown that these are oftentimes followed by abnormal stock returns (Muntermann and Güttler 2007).

We only include those corporate disclosures into our dataset that were published during stock exchange trading hours because we investigate intraday liquidity effects. As below event study analysis requires 15 minutes after / prior to ad hoc publication as input, the earliest time for inclusion into the dataset is 9:15 a.m. and the latest is 5:15 p.m. Moreover, we concentrate on those companies that were member in one of the following German stock indices at publication date: DAX (large-capitalization stocks), MDAX (medium-capitalization stocks), and SDAX (small-capitalization stocks). The final dataset comprises of 396 ad hoc disclosures published between 2006-01-31 and 2008-09-09.

The corporate disclosures' associated companies are, at least, traded on the fully-electronic trading system Xetra. For each security a (open) limit order book is provided on Xetra (see Figure 1). Investors post orders into the limit order book and thereby indicate their willingness to trade. The dataset at hand contains this high-frequency (level-10) order book data, i.e. it allows insights into order book depth. The respective order book data was extracted from *Reuters Tick History* (university access). Respective order book data was extracted for event dates, i.e. the publication of ad hoc disclosures, and previous ten working days.

3.2 Methodology

The liquidity measure used in this study is similar to the *Cost of Round Trip* (*CRT*) proposed by Irvine, Benston and Kandell (2000). *CRT* builds upon the information contained in the open limit order book (see Figure 1); i.e. it measures the ex ante committed liquidity immediately available in the market. We compute $CRT(d)_t$ as follows: Given a certain order book situation (and the prices / quantities respectively) at time *t*, the cost of simultaneously buying and selling *d* units is calculated. This cost figure is divided by the number of units, i.e. *d*, to receive a per-unit cost of a roundtrip trade. High values in $CRT(d)_t$ indicate that the market is illiquid and implicit transaction costs are expected large in size. Any order book activity such as the submission, cancellation, adjustment or execution of (limit) orders naturally changes the $CRT(d)_t$. As we operate with time periods rather than with single points in time, we need to calculate an $averageCRT(d)_{T,[t1,t2]}$ for a specific time interval [*t1*, *t2*] at day *T*. Thereby, the $CRT(d)_t$ values at fixed points in time, i.e. every ten seconds, serve as input.

	BID	ASK	
Quantity	Limit	Limit	Quantity
150	52.00	53.00	145
200	51.00	54.00	65
50	50.00	55.00	100
75	49.00	56.00	110
100	48.00	57.00	200
…	…	…	…

Figure 1. Limit order book

$$ACRT(d)_{[t1,t2]} = \frac{averageCRT(d)_{T0,[t1,t2]} - \frac{1}{N}\sum_{j=1}^{N} averageCRT(d)_{T0-j,[t1,t2]}}{\frac{1}{N}\sum_{j=1}^{N} averageCRT(d)_{T0-j,[t1,t2]}} \quad \textbf{(1)}$$

To investigate the potential impact from the publication of corporate disclosures on firm liquidity levels, we make use of event study methodology. An approach similar to the constant-mean-return model is applied (Campbell, Lo and MacKinlay 1997). Thereby, we adjust an observed effect by a mean that has been calculated using historical data prior to event date (see Formula 1). Calculating the abnormal liquidity measure $ACRT(d)_{[t1,t2]}$ we adjust the $averageCRT(d)_{T0,[t1,t2]}$ at event day *T0* by previous *N* days (*N* = 10) $averageCRT(d)_{T,[t1,t2]}$ for the same time period [*t1*, *t2*]. The adjustment is undertaken for the same (intraday) time period [*t1*, *t2*] because literature suggests that liquidity levels systematically change during the course of the day (McInish and Wood 1992). As abnormal liquidity levels should be comparable among companies, these are calculated as relative values.

3.3 Results

$ACRT(d)_{[t1,t2]}$ is calculated for different sizes *d* and different time periods *[t1, t2]* (Table 2). If the mean of $ACRT(d)_{[t1,t2]}$ turns out to be significantly different from zero, the chosen event type *regulatory-driven corporate disclosures* constitutes a critical market event for which significantly higher / lower liquidity levels and consequently transaction costs can be expected.

First, it can be observed that there are no significant abnormal liquidity levels prior to the publication of corporate disclosures ([-5, 0]). This finding provides evidence that the chosen event type actually contains new previously unknown information that has not been widely anticipated by market participants. It is therefore expected that forecasting of such liquidity shocks is especially challenging for models, which are solely based on quantitative data.

Second, we find strong empirical evidence that transaction costs increase subsequent to the publication of corporate disclosures. This is most likely due to the fact that the disclosures' contents persuade traders to adjust their company valuations and adjust their existing limit orders accordingly. During the adjustment process less limit orders remain in the market and therefore the cost of execution increases, i.e. liquidity decreases. The speed of adjustment, however, seems to be different among index members. While $ACRT(500)_{[0, 5]}$ is statistically significant only for those corporate disclosures associated with companies being member of the DAX, later *ACRT(500)*$_{[5,\ 10]}$ is significant only for those corporate disclosures associated with companies being member either of MDAX or SDAX. One possible explanation for this finding is that German blue-chip DAX companies are monitored by a larger group of traders and therefore the information dissemination process is faster. Further support for the *information dissemination* hypothesis is given by the fact that even *ACRT(500)*$_{[10,\ 15]}$ is significant for those disclosures associated with smallest-capitalization stocks.

		ACRT(500)$_{[t1, t2]}$				ACRT(2000)$_{[t1, t2]}$			
Index	n	[-5, 0]	[0, 5]	[5, 10]	[10, 15]	[-5, 0]	[0, 5]	[5, 10]	[10, 15]
DAX	113	0.05	0.63 ***	-0.17	-0.17	0.02	-2.00	-1.45	0.00
MDAX	116	0.01	2.22	0.41 ***	2.91	-0.05	2.07 *	0.33 ***	-0.03
SDAX	167	0.44	0.54	0.40 ***	0.51 ***	0.16	0.19	0.22 **	0.36

***** / ** / * indicate significance at the 1% / 5% / 10%-level.*

Table 2. Mean ACRT(d)$_{[t1,\ t2]}$ *sorted by indices*

In line with Graham et al. (2006), we suppose that the reaction of liquidity levels to new information depends on news types. We therefore construct news sub-samples according to the categorization proposed by Leis and Nowak (2001) (Table 3). The sub-sample analysis provides the following insights: First, corporate disclosures belonging to the category *financial statement* do not exhibit statistically significant abnormal liquidity levels. This might be due to the fact that the dates for the publication of quarterly / final financial results are usually known well in advance. Therefore, the corporate disclosure itself, not necessarily the content, may be anticipated by market participants. Second, strongest empirical evidence for statistically significant abnormal liquidity levels can be found for news category *miscellaneous*. This may be due to the fact that these news are particularly unanticipated by market participants. This finding is consistent with Graham et al. (2006) who conclude that market reactions differ depending on whether an event's timing is known beforehand.

		ACRT(500)$_{[t1, t2]}$				
News category	n	[0, 15]	[-5, 0]	[0, 5]	[5, 10]	[10, 15]
All news categories	396	0.67 ***	0.20	1.06 *	0.24	1.02
(1) Financial statement	114	0.24	-0.01	-0.45	-0.11	2.33
(2) Dividend announcement	28	0.80 **	0.12	0.88	0.29 **	0.76 *
(3) Corporate action	32	1.32 **	0.20	2.95 *	0.81 *	0.66
(4) M&A transaction / reorganization	80	1.38 *	0.01	3.72	0.35 ***	0.54 *
(5) Personnel	61	0.40 *	0.27	0.17 **	0.27	0.32
(6) Litigation [a]	4	0.41	0.15	0.47	0.47	0.31
(7) Order situation [a]	3	1.88	0.19	1.30	1.31	2.88
(8) Investments [a]	10	0.57	0.05	0.29	-0.24	0.11
(9) Miscellaneous	64	0.43 ***	0.80	0.53 ***	0.42 ***	0.34 ***

[a] *As the number of observations, i.e. n, is well below 30 no test statistics are provided.*
***** / ** / * indicate significance at the 1% / 5% / 10%-level.*

Table 3. Mean ACRT(500)$_{[t1,\ t2]}$ *sorted by news categories*

Given above results, the simplest strategy to avoid high implicit transaction costs would be to either execute orders immediately at t_0 (*naïve* strategy) or wait for execution until liquidity reverts to normal levels after approximately 10 to 30 minutes (see Figure 5). The latter case would, however, incur waiting / opportunity costs. We therefore conclude that a simple binary trading signal of the kind *news* or *no news* (incl. news category *financial statement*) might already be helpful for *automated traders*.

4 INTRADAY TEXT MINING APPROACH

The purpose of this section is to develop a model-based system that is able to automatically assess regulatory corporate disclosures' contents with regard to its' near-future influence on individual securities' liquidity levels. The system is intended to produce trading signals that serve *automated trading engines* as input. The proposed text mining approach is applied to above introduced dataset and evaluated by means of *classic* and domain-specific *simulation-based* model evaluation. An illustration of the setup can be found in Figure 4.

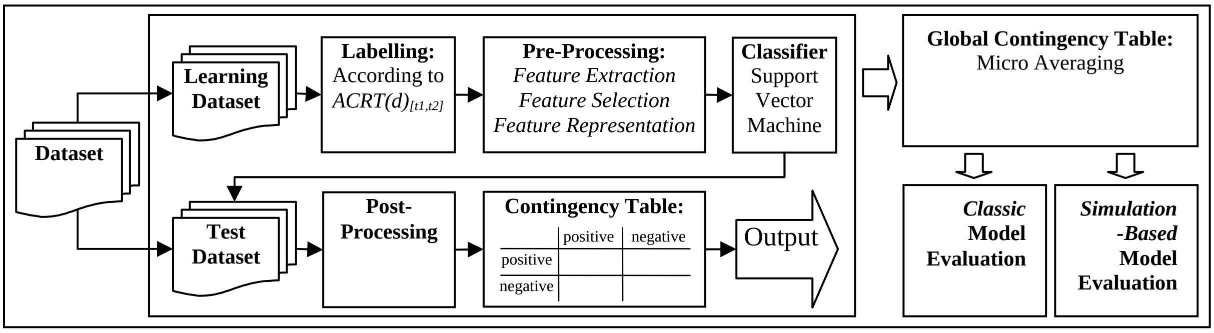

Figure 4. Intraday text mining approach setup

4.1 Labelling

Supervised learning requires the documents in the document collection to be labelled according to pre-defined objectives. Our objective is to forecast news-related liquidity impacts: As the news category *financial statement* does not reveal statistically significant abnormal liquidity levels, we exclude it from the following analysis (dataset). If this news category remained in the sample and if we tried to identify those $ACRT(500)_{[t1,\, t2]}$ well below zero, our text mining approach might be *tempted* to simply identify those corporate disclosures that belong to the news category *financial statement*. Such a category classifier, however, is not that valuable because the news categories are already part of the provided data stream (meta-data). For the remaining 282 corporate disclosures we can, on average, expect abnormal liquidity levels.

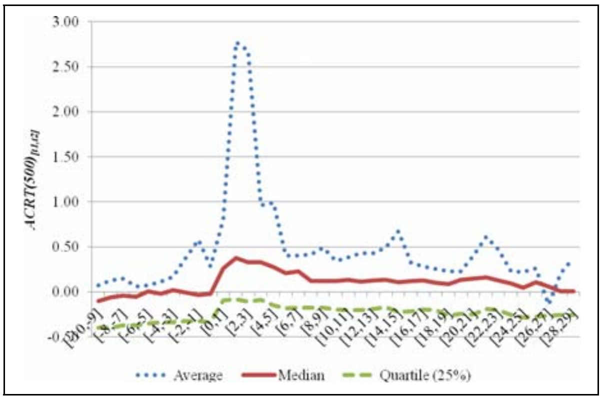

Figure 5. $ACRT(500)_{[t1,\, t2]}$ event study results

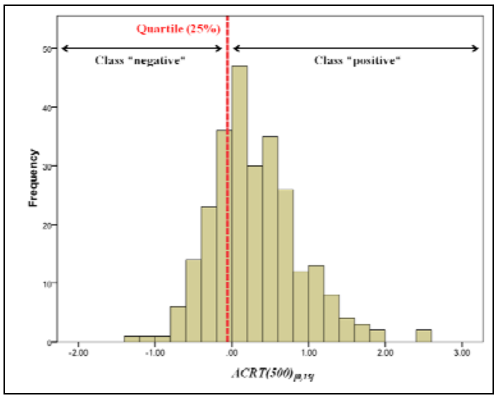

Figure 6. $ACRT(500)_{[0,\, 15]}$ distribution

The text mining task is to identify those corporate disclosures that are associated with negative $ACRT(500)_{[t1,\, t2]}$, i.e. those where transaction costs are still below historical levels. As shown in Figure 5 the $ACRT(500)_{[t1,\, t2]}$ 25% quartile stays below zero for all periods. Therefore, each corporate disclosure is assigned to the classes *positive* or *negative* depending on whether their $ACRT(500)_{[t1,\, t2]}$ is above or below the 25% quartile of all documents' $ACRT(500)_{[t1,\ t2]}$ (Figure 6). For labelling, we use the time period [0, 15] because it revealed to adequately summarize sub-period results (Table 3).

4.2 Text Pre-Processing

During pre-processing corporate disclosures are transformed into a numeric representation. This is necessary because *traditional* machine learning methods are not able to cope with plain text. In line with Brücher, Knolmayer and Mittermayer (2002) we employ the three pre-processing steps *feature extraction*, *feature selection*, and *feature representation*: During *feature extraction* a dictionary of words and phrases that describes the document collection adequately is generated. Thereby, a simple *StringTokenizer* (Wurst 2009) splits up the whole text into individual units. In order to eliminate *noise*, i.e. words with little meaning but frequent appearance, a stop word list and a threshold on the number of documents each token occurs in, has been made use of. Moreover, different grammatical forms of a word are mapped to a common stem by applying the *Porter Stemmer* (Porter 1980). We do not apply additional *feature selection* methods for the following reasons: First, Forman (2003) finds that the available feature selection methods do not necessarily "perform better than using all features available". Second, Joachims (1998) provides evidence that even features ranked lowest still contain considerable information and are somewhat relevant. Third, the applied classification method *SVM* is "nearly independent of the dimensionality of the feature space" (Hotho, Nurnberger and Paaß 2005) and is therefore not expected to suffer from the "curse of dimensionality". During *feature representation* each document is represented by previously extracted and selected number of features. The respective feature weightings in the document-feature matrix *W* is given by *tfidf* (Lewis 1992).

4.3 Classification Technique

We use Support Vector Machines (*SVM*) within our text mining approach. Comparative empirical studies provide evidence that the classification performance of *SVM* is superior to other data mining techniques (e.g. Joachims 1998). In addition *SVM* "is usually less vulnerable to the over-fitting problem [and] the solution of *SVM* is always unique and globally optimal" (Huang, Nakamori and Wang 2005). *SVM* was first introduced by Vapnik (1995) for solving two-class recognition problems. The basic idea is to find a decision surface that maximizes the margin between data points, i.e. classes, by means of structural risk minimization. In case of originally non-separable data points, the original data vectors may be mapped to higher dimensional space to achieve linear separability again. In order to reduce complexity, *kernels*, i.e. functions in lower dimensional space that exhibit similar behaviour as the original functions in higher dimensional space, are applied. We are making use of a linear *kernel* because Hsu, Chang and Lin (2003) provide evidence that a linear *kernel* seems sufficient whenever the number of features is exceptionally large.

4.4 Post-Processing

SVM delivers confidence values for each class as a kind of guarantee that the prediction is actually *true positive* (Mierswa et al. 2006). Corporate disclosures are assigned to the classes *positive* or *negative* depending on whether the confidence value is above or below a certain (learned) threshold. The variation of thresholds may be applied as a post-processing step to account for imbalanced datasets or unequal classification costs (Yang 2001). We are confronted with both: First, the class *negative*, which is of most interest to us, merely contains 25% of all corporate disclosures. Second, the cost for falsely classifying *negative* documents as *positive* is higher than falsely classifying *positive* documents as *negative*. This is because our objective is to precisely identify those corporate disclosures that are associated with abnormally low transaction costs after publication, i.e. class *negative*. To overcome these problems, we conduct cost-sensitive learning implemented as a post-processing step (Ikonomakis, Kotsiantis and Tampakas 2005). Thereby, a *ThresholdFinder* (Mierswa et al. 2006) uses the confidence values to turn the *SVM* into a cost-sensitive learner.

5 *CLASSIC* MODEL EVALUATION

5.1 Classic Model Evaluation Setup

As shown in Figure 4, the whole dataset is split up into a learning dataset and a test dataset in order to ensure that model evaluation is independent from model building. As the dataset is comparatively small we do not conduct a one-time split, but follow a *k-fold* (k = 10) cross validation approach (Witten and Frank 2001). Each test sub-sample contingency table is aggregated to create a global contingency table (micro averaging) (Table 7). The global contingency table is used to calculate the *classic* performance measures *accuracy* [(a+d)/n; n=a+b+c+d], *recall* [class *positive*: a/(a+c)], and *precision* [class *positive*: a/(a+b)] (Hotho, Nurnberger, and Paaß 2005).

	True *positive*	True *negative*
Predicted *positive*	*a*	*b*
Predicted *negative*	*c*	*d*

Table 7. Exemplary illustration of (global) contingency table

5.2 *Classic* Model Evaluation Results

Classification results for the class *negative* (Table 8) provide the following insights: High (misclassification) costs of the class *negative*, e.g. 0.9, result in a high *precision* figure and a low *recall* figure. In other words, there are only very few corporate disclosures assigned to the class *negative*, but those that are assigned actually truly belong to this class. This is due to the fact that corporate disclosures are merely assigned to the class *negative* if the respective *SVM* confidence value is quite high, i.e. the classifier is “certain” that the disclosure actually belongs to the respective class.

Misclassification cost for class		Accuracy	Recall [a]	Precision [a]
positive: 0.1	*negative*: 0.9	78.37 %	14.08 %	100 %
positive: 0.3	*negative*: 0.9	78.72 %	21.13 %	78.95 %
positive: 0.5	*negative*: 0.9	77.66 %	42.25 %	57.69 %
positive: 0.7	*negative*: 0.9	75.90 %	49.30 %	52.24 %
positive: 0.9	*negative*: 0.9	72.34 %	56.34 %	45.98 %
positive: 0.9	*negative*: 0.1	37.59 %	100 %	28.74 %

[a] *Classification results are provided for the class of most interest, i.e. negative.*

Table 8. Classic model evaluation results

The *precision* figure of 100% for $SVM_{(0.1;0.9)}$ provides strong evidence that the proposed text mining approach is able to precisely identify some of those corporate disclosures whose associated securities' liquidity levels are not negatively affected by the disclosures' contents. The high *precision* figure comes at the cost of low *recall*: We merely capture a small number of relevant class *negative* corporate disclosures. Therefore, one might be willing to accept a certain number of "*false positives*" to increase the number of caught class *negative* corporate disclosures, e.g. $SVM_{(0.5;0.9)}$. At the end, it is up to the developers of *automated trading engines* to decide upon the classified documents' confidence values. For instance, a *safe* trading signal would be provided at *precision* of 100% with $SVM_{(0.1;0.9)}$. *Accuracy* figures are slightly larger than the 75% all-*positive* benchmark would suggest. As we are, however, primarily interested in precisely identifying class *negative* corporate disclosures, the benchmark is of little use. We merely show *accuracy* figures for reasons of completeness.

To conclude, the proposed system is able to produce a trading signal that indicates whether or not the underlying regulatory corporate disclosures will (most likely) cause abnormally high or low transaction cost levels during the 15 minutes subsequent to their publication. This trading signal may serve as additional input into existing trading models purely based on quantitative data.

6 *SIMULATION-BASED* MODEL EVALUATION

6.1 *Simulation-Based* Model Evaluation Setup

Above text mining approach and the resulting trading / liquidity signals will be further evaluated by means of a simulation. Previous applications of text mining techniques in the financial industry have highlighted the need for domain-specific evaluation metrics (Groth et al. 2009). The *novel* simulation setup constitutes an *automated trading engine* use case. The simulation is, however, primarily intended to merely serve as an additional text mining classification evaluation. Therefore, certain model assumptions may not perfectly reflect reality: Within the simulation an *automated trader* receives the order to execute *S* shares during time interval *M [m1, m2]*. The goal is to minimize implicit transaction costs. In order to achieve this goal and derive an optimal execution strategy, the *automated trader* needs to forecast future liquidity levels. The decision in favour of one or the other trading strategy is *static* in our simulation model, i.e. the trading strategy chosen at *m1* cannot dynamically be altered during *M*. Moreover, we do not explicitly model buy or sell orders. Instead, each time the *automated trader* triggers an execution of size s_i, above defined $CRT(s_i)_{ti}$ is used as a proxy for incurred transaction costs. Transaction costs are aggregated according to Formula 3.

$$mp = \frac{1}{15}\sum_{i=0}^{15}\frac{bestBID_{t_i} + bestASK_{t_i}}{2} \qquad \textbf{(2)}$$

$$Cost_{Strategy} = \frac{\sum_{i=0}^{15} CRT\left(s_i\right)_{t_i}}{mp} \qquad \textbf{(3)}$$

The simulation time interval *M* is equivalent to the period used for labelling, i.e. [0, 15]. *S* = 1,600 shares need to be executed during this time interval in each security being subject to a corporate disclosure. Executions of size s_i are made at fixed one-minute intervals (t_0, t_1, t_2, …t_{15}). Thereby, t_0 constitutes the corporate disclosure publication time at publication day. It is assumed that *automated traders* – even with a simple trading signal as *news* or *no news* – are able to react to the publication of corporate disclosures within milliseconds, i.e. while liquidity is still at *normal* levels (at t_0).

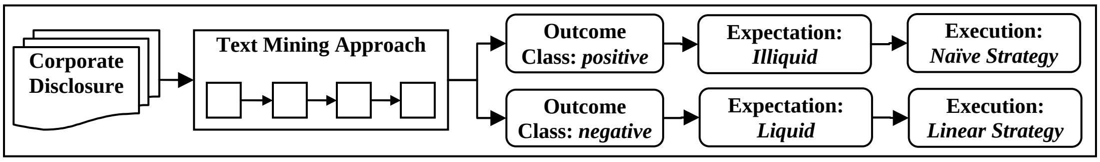

Figure 9. Setup of simulation-based model evaluation

Our *k*-fold cross validation simulation setup therefore looks as follows (see Figure 9): Each corporate disclosure in the *k* test datasets is classified by above proposed text mining approach. If the corporate disclosure is classified as belonging to the class *positive*, we expect the liquidity of the underlying security to decrease during time interval [0, 15]. Consequently, the *automated trader* should prefer to execute at the beginning of the interval. The execution of all shares at t_0 is called *naïve* strategy. It shall be noted that large trades may entail comparatively larger market impacts because of non-linear market impact functions. If the corporate disclosure is classified as belonging to the class *negative*, we expect the liquidity of the underlying security to increase during time interval [0, 15]. Consequently, the *automated trader* triggers a strategy that simultaneously entails a smaller market impact per trade and profits from generally higher / better liquidity levels. This strategy equally / *linearly* executes volume *S* over period *M*, i.e. s_i = 100 at t_0, t_1, t_2 etc. The *linear* strategy is merely accompanied by a temporary market impact that does not last until the next execution time t_i.

6.2 *Simulation-Based* Model Evaluation Results

Descriptive results for the *simulation-based* model evaluation and varying *SVM* misclassification cost inputs provide the following insights (see Table 10): First, if either the *naïve* strategy or the *linear* strategy is adhered to at each event (i.e. n = 282) irrespective of the disclosures' contents the *naïve* strategy reveals to be superior because of lower incurred transaction costs (1.85 vs. 2.67). It follows that the proposed binary *news / no news* trading signal may already help to save transaction costs. Second, results provide evidence that the text mining approach is able to precisely identify (some of) those corporate disclosures that do not entail negative liquidity shocks. For example, given the misclassification cost setting (0.1;0.9) the *SVM* model classified 10 corporate disclosures as *negative* and 272 corporate disclosures as *positive*. Given above described setup, the preferred strategy (see bold numbers) for events, i.e. news, classified as *negative* should be *linear*. As the *linear* strategy entails lower transaction costs than the alternative *naïve* strategy, i.e. -0.46 vs. 2.01, it seems as if the classifier correctly identified class *negative* corporate disclosures. Analogue, the preferred strategy for the corporate disclosures classified as *positive* is *naïve*. In these cases we can also observe lower transaction costs for the preferred strategy, e.g. 1.85 vs. 2.78. To conclude, the results provide first evidence that the *SVM* model correctly classified many class *positive* and class *negative* disclosures.

Classification inputs	Result subset actually classified as	n	$Cost_{Strategy}$ in % for execution strategy: *naïve*			*linear*		
			Mean	Median	Stdev.	Mean	Median	Stdev.
-	-	282	1.85	1.00	2.44	2.67	1.35	1.37
$SVM_{(0.1;\,0.9)}$	*negative*	10	2.01	1.63	1.69	**-0.46**	0.75	5.13
$SVM_{(0.1;\,0.9)}$	*positive*	272	**1.85**	1.31	2.47	2.78	1.85	6.02
$SVM_{(0.3;\,0.9)}$	*negative*	19	2.53	1.75	2.80	**-0.12**	1.41	5.62
$SVM_{(0.3;\,0.9)}$	*positive*	263	**1.80**	0.97	2.41	2.87	1.33	6.00
$SVM_{(0.5;\,0.9)}$	*negative*	52	2.03	1.22	2.16	**1.49**	1.10	5.03
$SVM_{(0.5;\,0.9)}$	*positive*	230	**1.81**	0.97	2.50	2.93	1.43	6.19

Table 10. Simulation-based model evaluation descriptive statistics ($Cost_{Strategy}$)

To further statistically explore this, corresponding hypotheses are formulated and tested (Table 11).

For subset *negative*:
$H_1 : \mu\left(Cost_{SVM[Strategy:linear]}\right) \geq \mu\left(Cost_{SVM[Strategy:naive]}\right)$

For subset *positive*:
$H_2 : \mu\left(Cost_{SVM[Strategy:naive]}\right) \geq \mu\left(Cost_{SVM[Strategy:linear]}\right)$

If a null hypothesis can be rejected, we can statistically corroborate a higher population mean of one execution strategy compared to another execution strategy for that classification cost setting at a given level of significance. We conduct two-sample t-tests with a hypothesized mean difference of zero.

Null Hypothesis	Data subset	*df*	t-value
$\mu\left(Cost_{SVM_{(0.1;0.9)}[Strategy:linear]}\right) \geq \mu\left(Cost_{SVM_{(0.1;0.9)}[Strategy:naive]}\right)$	*negative*	18	1.44 *
$\mu\left(Cost_{SVM_{(0.3;0.9)}[Strategy:linear]}\right) \geq \mu\left(Cost_{SVM_{(0.3;0.9)}[Strategy:naive]}\right)$	*negative*	36	1.84 **
$\mu\left(Cost_{SVM_{(0.5;0.9)}[Strategy:linear]}\right) \geq \mu\left(Cost_{SVM_{(0.5;0.9)}[Strategy:naive]}\right)$	*negative*	52	0.71
$\mu\left(Cost_{SVM_{(0.1;0.9)}[Strategy:naive]}\right) \geq \mu\left(Cost_{SVM_{(0.1;0.9)}[Strategy:linear]}\right)$	*positive*	542	2.37 ***
$\mu\left(Cost_{SVM_{(0.3;0.9)}[Strategy:naive]}\right) \geq \mu\left(Cost_{SVM_{(0.3;0.9)}[Strategy:linear]}\right)$	*positive*	524	2.67 ***
$\mu\left(Cost_{SVM_{(0.5;0.9)}[Strategy:naive]}\right) \geq \mu\left(Cost_{SVM_{(0.5;0.9)}[Strategy:linear]}\right)$	*positive*	458	2.55 ***

***** / ** / * indicate significance at the 1% / 5% / 10%-level.*

Table 11. Simulation-based model evaluation hypotheses testing

The null hypotheses can be rejected at high levels of significance for those cases with high class *negative* misclassification costs. In other words, the preferred strategies (see Figure 9) entail significantly lower transaction costs than the alternative strategies on the same data subsets. To conclude, the text mining approach produces a viable trading signal on expected future liquidity levels.

7 CONCLUSION

Liquidity constitutes one of the most important determinants of (implicit) transaction costs. We show by means of an empirical event study that the publication of regulatory corporate disclosures is followed by abnormal liquidity levels. This finding is consistent with existing beliefs on limit order traders updating their limit orders upon the arrival of new information. We, however, do not find evidence of abnormal liquidity levels prior to the publication of corporate disclosures. It follows that *automated traders* should ideally include information on the publication of corporate disclosures into their models, but these are not able to do this sufficiently fast based on purely quantitative data. Against this background, we propose two trading signals to be integrated into existing models:

(1) Due to the fact that the majority of corporate disclosures are associated with abnormally high transaction costs subsequent to their publication, a rather simple binary trading signal is proposed. The signal indicates whenever a corporate regulatory disclosure – except on the category *financial statement* – is published. *Automated traders* may then expect abnormal liquidity levels for this security and alter their execution strategies accordingly. The results of a *simulation-based* evaluation use case provide first evidence that the simple binary *news / no news* signal constitutes a viable trading signal.

(2) Due to the fact that some corporate disclosures are associated with lower transaction costs subsequent to their publication than others, the second trading signal marks these. The trading signal is based on a sophisticated text mining approach that automatically interprets the contents of the respective corporate disclosures. Both *classic* and *simulation-based* model evaluation results provide evidence that the trading signal precisely indicates some of those corporate disclosures entailing lowest expected future transaction costs. In other words, the second trading signal provides detailed insights for the first trading signal’s *news* (publication) case.

To conclude, we have proposed and successfully tested different ways on how to enhance *automated trading engines* to cope with news-related liquidity shocks in a timely manner. Future work will mainly concentrate on solving current limitations of research: First, the dataset shall be extended. Second, the proposed forecasting approach shall be compared to existing quantitative forecasting approaches. Please, however, note that this paper’s proposed trading signals are not intended to replace existing models, but shall rather complement them. Future work will therefore also concentrate on the integration of these trading signals into existing execution models.

References

Bikker, J., Spierdijk, L., Hoevenaars, R. and van der Sluis, P.J. (2006). Forecasting Market Impact Costs and Identifying Expensive Trades. Working Paper, University of Twente, Netherlands.

Brooks, R., Patel, A. and Su, T. (2003). How the Equity Market Responds to Unanticipated Events. Journal of Business, 76, 109-133.

Brücher, H., Knolmayer, G., and Mittermayer, M.-A. (2002). Document Classification Methods for Organizing Explicit Knowledge. In European Conference on Organizational Knowledge, Learning and Capabilities, Athens, Greece.

Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997). The Econometrics of Financial Markets. Princeton University Press, Princeton, USA.

Coggins, R., Lim, M. and Lo, K. (2006). Algorithmic Trade Execution and Market Impact. Working Paper, University of Sydney.

Deutsche Börse AG (2008). Integrity for Financial Markets. Annual Report 2008. http://www.deutsche-boerse.com, [2009-08-09].

Domowitz, I. and Yegerman, H. (2005). The Cost of Algorithmic Trading – A First Look at Comparative Performance (Bruce, B.R. Ed.), pp. 30-40, Algorithmic Trading: Precision, Control, Execution. Institutional Investor Inc.

Forman, G. (2003). An Extensive Empirical Study of Feature Selection Metrics for Text Classification. The Journal of Machine Learning Research, 3, 1289-1305.

Gomber, P., Schweickert, U. and Theissen, E. (2004). Zooming in on Liquidity. In 31st Annual Meeting of the European Finance Association, Maastricht, Netherlands.

Gomber, P. and Gsell, M. (2006). Catching Up with Technology – The Impact of Regulatory Changes on ECNs/MTFs. Competition and Regulation in Network Industries, 1 (4), 535-557.

Graham, J., Koski, J.L. and Loewenstein, U. (2006). Information Flow and Liquidity Around Anticipated and Unanticipated Dividend Announcements. Journal of Business, 79 (5), 2301-2336.

Groth, S.S. and Muntermann, J. (2009). Supporting Investment Management Processes with Machine Learning Techniques. In 9. Internationale Tagung Wirtschaftsinformatik, Vienna, Austria, 275-284.

Hotho, A., Nurnberger, A. and Paaß, G. (2005). A Brief Survey of Text Mining. GLDV Journal for Computational Linguistics and Language Technology, 20 (1), 19-62.

Hsu, C.-W., Chang, C.-C. and Lin, C.-J. (2003). A Practical Guide to Support Vector Classification. Technical Report, National Taiwan University, Taiwan.

Huang, W., Nakamori, Y. and Wang, S.Y. (2005). Forecasting Stock Market Movement Direction With Support Vector Machine. Computers and Operations Research, 32 (10), 2513-2522.

Ikonomakis, M., Kotsiantis, S. and Tampakas, V. (2005). Text Classification Using Machine Learning Techniques. WSEAS Transactions on Computers, 4 (8), 966-974.

Irvine, P., Benston, G. and Kandel, E. (2000). Liquidity Beyond the Inside Spread: Measuring and Using Information in the Limit Order Book. Working Paper, Emory & Hebrew University .

Joachims, T. (1998). Text Categorization with Support Vector Machines: Learning with Many Relevant Features. In 10th European Conference on Machine Learning, Chemnitz, 137-142.

Krinsky, I. and Lee, J. (1996). Earnings Announcements and the Components of the Bid-Ask Spread. Journal of Finance, 51, 1523-1535.

Lee, Ch., Mucklow, B. and Ready, M. (1993). Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis. Review of Financial Studies, 6, 345-374.

Leis, J. and Nowak, E. (2001). Ad-hoc-Publizität nach 15 WpHG. Schäffer-Poeschel, Stuttgart.

Lewis, D.D. (1992). Representation and Learning in Information Retrieval. Dissertation, University of Massachusetts, Amherst, USA.

McInish, T.H. and Wood, R.A. (1992). An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks. Journal of Finance, 47 (2), 753-764.

Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M. and Euler, T. (2006). YALE: Rapid Prototyping for Complex Data Mining Tasks. In ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Philadelphia, USA.

Mittermayer, M.-A. and Knolmayer, G.F. (2006). Text Mining Systems for Market Response to News: A Survey. Working Paper No. 184, University of Bern.

Muntermann, J. and Güttler, A. (2007). Intraday Stock Price Effects of Ad Hoc Disclosures: The German Case. Journal of International Financial Markets, Institutions and Money, 17 (1), 1-24.

Porter, M.F. (1980). An Algorithm for Suffix Stripping. Program, 14 (3), 1980, 211-218.

Ranaldo, A. (2003). Intraday Market Dynamics Around Public Information Arrivals. Working Paper.

Vapnik, V. (1995). The Nature of Statistical Learning Theory. Springer, New York, USA.

Witten, I. H. and Frank, E. (2001). Data Mining: Praktische Werkzeuge und Techniken für das Maschinelle Lernen. Hanser Verlag, Munich, Vienna.

Wuthrich, B., Cho, V., Leung, S., Permunetilleke, D., Sankaran, K. and Zhang, J. (1998). Daily Stock Market Forecast from Textual Web Data. In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, San Diego, USA.

Wurst, M. (2009). The Word Vector Tool: User Guide, Operator Reference, Developer Tutorial. http://wvtool.sf.net, [2009-05-21].

Yang, Y. (2001). A Study of Thresholding Strategies for Text Categorization. In Proceedings of the 24th ACM Int. Conference on Research & Development in Information Retrieval, USA, 137-145.