



How to enable automated trading engines to cope with news-related liquidity shocks? Extracting signals from unstructured data



Sven S. Groth, Michael Siering*, Peter Gomber

Goethe University Frankfurt, Grüneburgplatz 1, 60323 Frankfurt, Germany

ARTICLE INFO

Article history:

Received 9 July 2012

Received in revised form 10 February 2014

Accepted 6 March 2014

Available online 16 March 2014

Keywords:

Automated trading

Liquidity

Forecasting

Text mining

e-Finance

Simulation

ABSTRACT

Financial markets are characterised by high levels of complexity and non-linearity. Information systems have often been applied to support investors by forecasting price changes in securities markets. In addition to the asset price, liquidity represents another financial variable that has a high relevance for investors because it constitutes a main determinant of total transaction costs. Previous research has shown that the level of liquidity is affected by the publication of corporate disclosures. To derive an optimal order execution strategy that minimises the transaction costs, investors as well as automated trading engines must be able to anticipate changes in the available market liquidity. However, there is no research on how to forecast the impact of corporate disclosures on market liquidity. Therefore, we propose an IT artefact that allows automated trading engines to appropriately react to news-related liquidity shocks. The system indicates whether the publication of a regulatory corporate disclosure will be followed by a positive liquidity shock, i.e., lower transaction costs compared to historical levels. Utilising text mining techniques, the content of the corporate disclosures is analysed to generate a trading signal. Furthermore, the trading signal is evaluated within a simulation-based use case that considers English and German corporate disclosures and is shown to be of economic value.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Decision-making in the financial domain represents a challenging task because financial markets are characterised by high levels of complexity and non-linearity [8]. These market characteristics make it difficult for decision-makers to quickly adjust their strategies in cases of company-related or economic events. Here, information systems play a crucial role in supporting human and computer-based decision-makers alike.

Today, the group of computer-based *automated*¹ traders generates already approximately one-half of the trading activity on major European markets such as Deutsche Börse's Xetra, and the percentage share continues to grow [7]. Computer-based *automated traders* "emulate a broker's core competence of slicing a large order into a multiplicity of smaller orders and of timing these to minimise the market impact" [13]. The decision on the investment or portfolio allocation itself is performed by the respective portfolio manager at a fund management company, and the primary task of *automated traders* is to execute the orders that are received from these fund management companies or institutional investors at the best available conditions.

To determine an optimal execution strategy for a pre-defined execution time period, i.e., to achieve the best available conditions, *automated traders* must handle the trade-off between the evolving transaction cost components in the order's execution. Costs that are generated while implementing investment decisions can generally be divided into two broad categories: First, there are explicit costs, such as commissions, fees, and taxes. Second, there are implicit costs, such as market impact, timing costs, and opportunity costs [1]. Especially for large trades, implicit transaction costs are mostly much larger than explicit transaction costs. Liquidity constitutes the main determinant of implicit transaction costs: if the number of shares that other market participants are willing to trade at a given limit is reduced, then the market impact of an order is increased [9]. Thus, liquidity and implicit transaction costs represent two sides of the same coin [41]: The higher the liquidity is, the lower the implicit transaction costs, and vice versa. Thus, "asset managers and ordinary investors care about liquidity insofar as it affects the return on their investments, simply because illiquid securities cost more to buy, and sell for less" [12]. Therefore, to derive an optimal execution strategy and to minimise the associated transaction costs, the ability to forecast future liquidity levels is very important.

Bikker et al. [1], however, conclude that "forecasting market impact costs appears notoriously difficult and traditional methods fail". Moreover, Domowitz and Yegerman [10] find that the execution quality of *automated traders* is inferior to the executions that are handled by brokers. One possible reason for this observation might be the fact that the currently employed models are solely based on purely

* Corresponding author. Tel.: +49 69 798 33832; fax: +49 69 798 35007.

E-mail addresses: sgroth@wiwi.uni-frankfurt.de (S.S. Groth), siering@wiwi.uni-frankfurt.de (M. Siering), gomber@wiwi.uni-frankfurt.de (P. Gomber).

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

quantitative data input. Existing (academic) models and strategies largely neglect one of the most important sources of information, which is unstructured qualitative data (i.e., news) [6]. If, for example, a listed company issues an unanticipated regulatory ad hoc disclosure, then *automated traders* cannot react sufficiently fast simply because they cannot analyse the content. Because unanticipated news, by definition, is very unlikely to be reflected in quantitative time series data prior to its publication date, *automated traders* can respond only to other (human) market participants' reactions.

Against this background, our research goal is to investigate whether and how unstructured qualitative data can be used as input for *automated trading engines*. We are especially interested in whether useful information can be extracted automatically from qualitative data to predict future levels of liquidity after the publication of corporate disclosures. For extracting such information, text mining techniques are utilised. Given that decision support in the financial domain is very challenging [8], we especially call the reader's attention to those domain-specific issues that require an adjustment of standard knowledge discovery approaches, to investigate and emphasise the economic relevance of the proposed system. We particularise the knowledge discovery in databases (KDD) process proposed by Fayyad et al. [11] by domain-specific customisations, such as the application of an event study for *data understanding* and a novel evaluation scenario in the form of a trading simulation. Moreover, we investigate the role of language within the proposed text mining framework. In other words, we enquire whether the proposed text mining system is sensitive to the language of the input text.

Thus, we contribute to the literature on financial text mining by proposing an IT artefact to forecast the liquidity impact of corporate disclosures. In contrast to previous studies that mainly focus on forecasting the stock price impact of financial news, we concentrate on the most important criterion of market quality, which influences the highest cost component in trading, i.e., the implicit transaction costs [25]. Furthermore, we enhance previous research by focusing on the economic relevance of financial text mining systems by extending the KDD process by Fayyad et al. [11] by means of an event study and a novel simulation. In this way, the simulation aims at evaluating the economic value of the proposed system and extends previous studies by accounting for the timing of the orders. Finally, we investigate whether the results differ for different language inputs.

The remainder of this paper is structured as follows: First, we present related financial and information systems research. Second, we describe the study setup, which is based on the KDD process proposed by Fayyad et al. [11]. Third, the adjusted KDD process is applied to the above-described *automated traders'* use case. In this way, the dataset that is used in this study is described, and the liquidity impact of the publication of regulatory corporate disclosures is investigated, to enhance the data understanding. Fourth, building on these insights, we propose a text mining approach that predicts the liquidity impact of the ad hoc news. The classification quality is evaluated with respect to both the *classic* model evaluation metrics and domain-specific *simulation-based* model evaluation. Finally, we present our conclusions.

2. Related work

2.1. Financial text mining

There are several studies that apply text mining techniques in financial markets to find patterns in text that can serve for predictions. In this context, Mittermayer and Knolmayer [35] provide a good survey on existing text mining systems: Since the time that Wuthrich et al. [46] proposed one of the first financial text mining applications, systems have been 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), and novel

evaluation methods [17]. 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 utilising text mining techniques. Additionally, the language of input texts has been addressed in the literature [3] but not with regard to the financial industry. Loughran and McDonald [31], however, have highlighted that financial texts require domain-specific knowledge and interpretation.

As a consequence, our approach builds upon and extends this literature by using intraday high-frequency data on the new forecasting object *liquidity*. We additionally concentrate on a certain news type and develop a novel domain-specific evaluation metric.

2.2. Liquidity impact of corporate disclosures

Liquidity refers to the possibility of buying or selling an asset immediately without adversely affecting the price, and it is seen to be the most important aspect of market quality [14]. Liquidity is composed of two key dimensions that are relevant in our context: A market has *breadth* when the best buy and sell orders exist in substantial volume (see Fig. 1). Additionally, a market has *depth* when there are orders in substantial volume in the closest neighbourhood to the best bid and best ask limits [41]. To evaluate breadth and depth, an open limit order book that displays the cumulated buy and sell orders can be used. Two exemplary open limit order books are shown in Fig. 1. On the left of each order book, the buy orders (bid) are displayed, whereas on the right, the sell orders (ask) are indicated; both have order limits and available quantity at their respective limits.

In general, a lower level of market breadth or market depth (and, consequently, a lower level of liquidity) is disadvantageous for market participants because these levels would implicate higher implicit transaction costs. Referring to the order book examples that are displayed above, an investor who is willing to buy a quantity of 145 stocks would have to pay 53.00 for each stock in order book situation 1. However, if there were only 100 shares available at the best ask of 53.00, the investor would have to buy additional 45 shares for 54.00 to obtain the desired quantity (order book situation 2). The resulting average price of 53.31 per share would be worse than it has been in the more liquid market. As a result, in order book situation 2, market participants would have to bear an increased amount of implicit transaction costs.

To summarise, liquidity refers to the bids and offers that are provided in the market and are listed in the order book, i.e., it is shown *ex-ante* which trading opportunities are available for traders. In contrast, prices and trading volume result from the liquidity demand, i.e., they reflect past trades, and because they are *ex-post*, they are not relevant for decisions that are related to the timing of orders. Therefore, in contrast to existing research, we focus on the relevant pre-trade decision criterion (i.e., the liquidity) instead of focusing on the impact of corporate disclosures on prices *ex-post*.

In previous studies, the effect of information arrival on company-specific liquidity levels has already been analysed (e.g., [28]). Most of the contributions, however, merely concentrate on one event type, such as earnings announcements [26] or dividend announcements [15]. Moreover, authors use liquidity measures that do not allow a complete analysis of transaction costs because they account only for the market breadth [39]. Finally, other studies are based on very few events and/or short time periods [4,14]. We address all of the above shortcomings in this paper to extend the previous literature: We analyse a comparatively large dataset that is composed 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 liquidity (i.e., the implicit transaction costs). Finally, we focus on short-term intraday liquidity effects.

Download English Version:

<https://daneshyari.com/en/article/554729>

Download Persian Version:

<https://daneshyari.com/article/554729>

[Daneshyari.com](https://daneshyari.com)