Market Impact Measurement of a VWAP Trading Algorithm

Jan Fraenkle

Svetlozar T. Rachev

Christian Scherrer

Jan Fraenkle
Statistics, Econometrics and Mathematical Finance, School of Economics and Business Engineering, KIT (Karlsruhe Institute of Technology), and Quantitative Analyst (Trading Algorithms), Phi-T Products & Services GmbH & Co KG
Email: jan.fraenkle@phi-tps.de

Svetlozar T. Rachev
Chair of Statistics, Econometrics and Mathematical Finance, School of Economics and Business Engineering, KIT, and Department of Statistics and Applied Probability, University of California, Santa Barbara, and Chief-Scientist, FinAnalytica INC Kollegium am Schloss, Bau II, 20.12, R210, Postfach 6980, D-76128, Karlsruhe, Germany
Email: zari.rachev@kit.edu

Christian Scherrer
Statistics, Econometrics and Mathematical Finance, School of Economics and Business Engineering, KIT, and Quantitative Analyst, Phi-T GmbH, Signal Provider for Lupus alpha NeuroBayes Fund.
Email: scherrer@phi-t.de

Abstract

We propose a model for market impact of algorithmic trades. Usually large orders cannot be executed immediately without significant trading costs. For optimized execution one relies on the help of a VWAP trading algorithm. We demonstrate that this VWAP algorithm is the optimal solution of the optimization problem using the presented market impact models.
The purpose of this work is the empirical market impact analysis of a homogeneous set of algorithmic trades. The underlying data set contains trades resulting form a hedge fund trading strategy. We found the participation rate to be the most important description variable. Therefore we provide a linear model and also a concave power law model of the market impact dependent on the participation rate. The estimated parameters lead to interesting consequences to verifying certain aspects of the market microstructure theory.
Our results also suggest different behavior of the various analyzed markets. On the one hand the market impact dependency on the participation rate behaves different for the Japanese market compared to the European, US, and Canadian. On the other hand, the individualized linear regression results suggest a dependency of the market impact on the tick size for the Japanese and the US market.
Keywords: market impact, trading costs, algo trading, trading volume, participation rate

1 Introduction

The performance of mutual funds strongly depends on transaction costs. For high frequency hedge fund strategies with a large turnover transaction costs thus play a crucial role. Very often the size of a fund is limited because too large sizes cannot be traded profitably: given the price predictions usually the market impact increases when trading volume becomes larger thus reducing the benefit of the strategy.

Transaction costs generally consist of two components: explicit costs including exchange and broker fees and also implicit costs such as market impact. Market impact is the interaction of a market participant's own activity on the market. In general the price observed at the beginning of a large trade is not equal to the actual execution price - on average the execution price is worse. This effect is also well known as implementation shortfall and discussed for example by Demsetz (1968) and Perold (1988).

The current work is an empirical analysis of the market impact of a homogeneous set of algorithmic trades in the stock market from April 2008 to July 2010 on Canadian, European, Japanese, and US stocks. It is done with the help of a proprietary data set originating from real trading activity of the Lupus alpha NeuroBayes®Short Term Trading Fund. We emphasize the remarkable size of this data set: more than 2 years trading activity on over 800 stocks in various countries (Europe, USA, Canada, and Japan) and a trade volume of over 30 billion USD equivalent. The main advantage of this data set compared to publicly available ones is due to the fact that single orders of one market participant are identified and connected. Because of that, the resulting market impact of the entire transaction can be measured. Without connecting single orders, the orders usually are assumed to be independent which is obviously not true. In that case, a trading strategy of a market participant cannot be reconstructed. So the characteristics of the proprietary data set enables us to provide rare empirical measurements to verify theoretical considerations.

The used trading algorithm is a so called VWAP (short for volume weighted average price) trading algorithm combined with a smart order router [1]. Its objective is the

[1] The purpose of the smart order router is to find the best trading venue in consideration of low exchange fees and a good execution price.

execution of a macro order within a given time at minimal execution cost. It is realized by splitting up the size of the macro order over the trading period according to the entire trading volume profile. We show that the VWAP trading algorithm is the strategy producing lowest market impact when taking the current market impact models as a basis.
The current work is an approach motivated by the needs of a practitioner. It provides several models describing the dependence between the market impact and some description variables. We found the participation rate as the by far most important variable to describe the market impact. As participation rate we define the ratio between macro order size and the entire number of traded shares in the respective period of time. The first market impact model is linear in the participation rate whereas the second one is based on a power law. We improve the explanation power of the linear model with the introduction of an individualized linear regression. In doing so, the linear regression parameters (slope and intercept) are dependent on additional exogenous variables. For this purpose we took the linear model: when we take this route for a portfolio optimization, it is much easier to find an optimum in a multidimensional space since the model for relative transaction costs is linear in order size. In contrast to arbitrary functions there are for linear and quadratic optimization functions very efficient and fast optimization algorithms, such as Simplex and Gauss-Newton.
The model based on a power law is motivated by the results of the microstructure theory and by a slightly concave curve observed in the data especially for a wide range of participation rates.
The different markets behave quite similar in many aspects, but we find significant differences between the Japanese market and the remaining markets. This can partly be explained by different regulations (such as the up-tick-rule in the Japanese market and different tick size definitions).
Additionally, we can show that the VWAP trading strategy is the optimal execution strategy taking the results of our market impact models into account.
The paper is organized as follows. In section 2 the data set and the trading algorithm is specified. Section 3 provides the market impact models with the participation rate as the only description variable. An individualized model, where the dependency of

several description variables and the market impact is discussed in section 4. It is shown in section 5 that the VWAP execution strategy is optimal. Section 6 concludes the paper.

2 Description of the Data Set

2.1 Trading Specifications

The underlying data sample of the current work contains all relevant informations about a set of about 120.000 algorithmic trades from April 2008 to July 2010. The data originates from the trading activity of the Lupus alpha NeuroBayes®Short Term Trading Fund [2]. Its stock universe consists of large caps from Europe, USA, Japan and Canada. More precisely, the universe consists of the 500 most liquid stocks in the USA, they are mainly covered by the S&P500 index. The 250 most liquid European stocks belong to the universe and also the 110 most liquid Japanese stocks and about 200 Canadian stocks. The investment strategy is based on statistical arbitrage on a day to day basis.

The implementation of the investment strategy, is realized with help of a trading algorithm. For trade execution the orders are split over a given time period and executed incrementally, since large orders cannot be executed at once at an attractive price (due to finite liquidity in the order book). The trading algorithm works on the basis of a VWAP trading strategy. This means the trading volume of the algo order is distributed over time, weighted by the entire trading volume, further details can be found in Fränkle and Rachev (2009).

All analyzed trades have the same trading period. In the US and Canadian market they take place in the last 15 minutes of the official trading hours, i.e. from 3.45 pm to 4.00 pm ET. The trades in the Japanese market are executed in the last 25 minutes of the trading day. The trading of European stocks is entirely different in the sense that there are closing auctions with significant trading volumes. So quite a large fraction of the algo orders is executed inside the closing auction. However, the execution of the algo order starts during the continuous trading session about 20

[2] http://www.ise.ie, Sedol: B1HMBP7

minutes before the closing auction.
The used trading algorithm tries to execute as much as possible with the usage of passive limit orders in order to reduce the market impact and explicit transaction costs.

2.2 Market Impact Definition

We define market impact as the interaction of the investors own order with the market, i.e. how large is the price change influenced by this order. Therefore the quantity market impact is described by a price difference between a benchmark price, which should be as little as possible influenced by the own order, and a price incorporating the full impact. In the current work, we take the relative price change r_s between the average execution price P_{vwap} and the arrival price $P_{arrival}$

$$r_s = d \cdot \frac{P_{vwap} - P_{arrival}}{P_{arrival}} \tag{1}$$

where

$$P_{vwap} = \frac{\sum_i p_i * q_i}{\sum_i q_i} \tag{2}$$

The direction d is 1 for all buys and -1 for the sells, p_i is the execution price of the i-th partial fill and q_i is the corresponding size. The arrival price $P_{arrival}$ is the valid stock price ultimately before the order arrives at the exchange. The average execution price is the volume weighted average price (vwap) of all transactions of the specific order during the trading period.
Note, however, that this relative price difference contains also the external triggered price movements which are not part of the market impact of the own order. So the price change r_s can be written as a sum of two components:

$$r_s = r_e + I \tag{3}$$

As already mentioned, one component is the stock price move r_e which is induced by external influences. The second component is the market impact I of the own trade which is analyzed here . To get rid of this effect we modify the market impact

definition in equation 1 as the mean value of the distribution of r_s. The advantage of this definition is that the mean value of the r_e distribution is 0, so the mean value of the r_s distribution is an unbiased estimator for the empirical market impact. The reason why this assumption is reasonable is market neutrality (dollar and beta [3] neutral) of the fund's investment strategy. Therefore the trades are also market neutral and market movements do not affect $< r_s >$[4].
For these reasons we can define the market impact of a trade as

$$< r_s >=< r_e > + < I >=< I > \tag{4}$$

without having a bias in the data by a nonzero mean of r_e. Although the external induced return does not contribute to the average impact $< I >$, it dramatically increases the variance of r_s. Therefore the width of the distribution of r_s is dominated by the external induced price movements. Hence the market impact can only be measured significantly with enough statistics.
The r_s distributions for the different markets are shown in the figures 1 and 2 together with the statistics in table 1. It is quite remarkable that the distribution of the transactions in the Japanese market has a huge peak at 0. This peak is explained by large tick sizes for many stocks, leading to a higher probability of unchanged stock prices. The average tick size over the execution price measured for the European universe is 4.7 BPS, for the US stocks 3.7 BPS and for the Japanese ons 18.0 BPS.

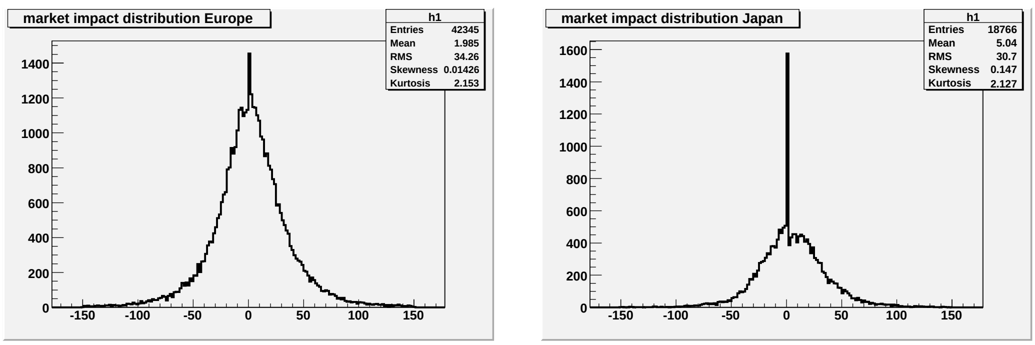

Figure 1: r_s distribution of Europe and Japan

[3]beta factor, known from the CAPM (Capital Asset Pricing Model)
[4]$< x >$ mean value of a set of numbers $x_1, x_2, x_3, ..., x_n$

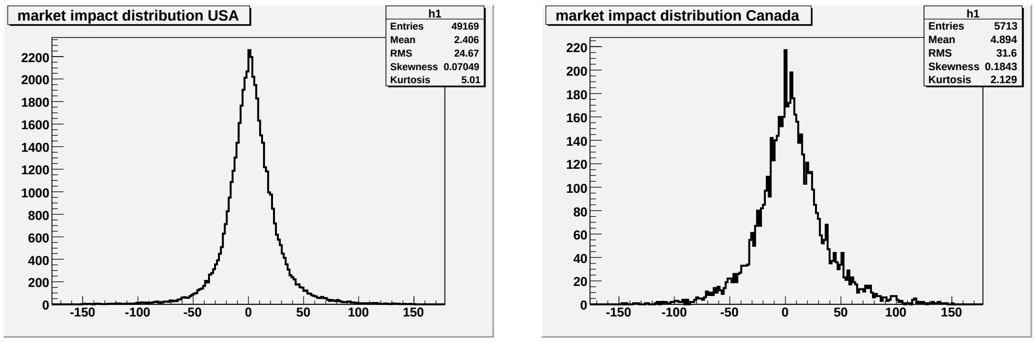

Figure 2: r_s distribution of USA and Canada

Table 1: Statistics of r_s distributions

	# trades	mean	error of mean	RMS
Europe	42056	2.006	0.17	34.35
Japan	18766	5.04	0.22	30.7
USA	48781	2.4	0.11	24.69
Canada	5493	4.84	0.43	31.66

2.3 Comparison of Sell and Short-Sell Trades

For all sells in the Japanese and US market the data provides the information whether the order was a long-sell or a short-sell. As explained in section 2.2 there is no bias in the average impact because $< r_e >= 0$ in equation 4. This is not the case if we take into account only subsets of the trades such as buy or sell orders. Still it is reasonable to look at the difference between sell orders and short-sell orders. Although we have a nonzero $< r_e^{\text{(short-)sell}} >$ as in figure we can approximate the expected return of the market to be the same for sell orders and short-sell orders:

$$< r_e^{\text{sell}} > \approx < r_e^{\text{short-sell}} > \tag{5}$$

Figure 3 shows the impact of trades in the Japanese and the US market for long-sell and short-sell orders. The US equities' impact distribution is not statistically significantly different for long-sells and short-sells because mean, standard deviation, skewness and kurtosis of both distributions cannot be distinguished by statistical tests. This is not the case for the trades in the Japanese market where the means of

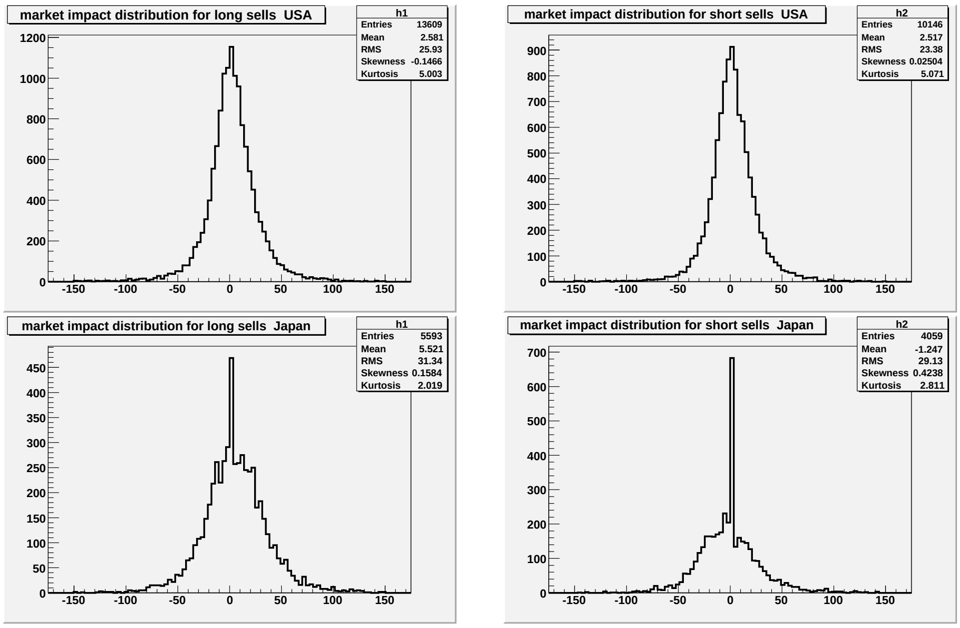

Figure 3: Comparison between the impact of normal sells and short-sells.

both distributions are significantly different [5]. The difference between the Japanese and the US market is due to a Japanese market rule, the so called 'up-tick rule'. For the rule to be satisfied, the short must be either at a price above the last traded price of the security, or at the last traded price if that price was higher than the price in the previous trade. So there is a bias in the execution of short-sells in getting better prices for short-sells. This comes for the price of a lower execution probability. Therefore the approximation 5 is not valid for Japan.

3 Market Impact Model

We analyze the dependency between the observable description variables and the market impact. As the most important variable we found the participation rate. It is defined as the ratio between the size of the algo order, of which the market impact is measured, and the entire number of shares of the security traded in the same time

[5] The statistical error of the mean σ_μ is defined by $\sigma_\mu = \sigma/\sqrt{n}$ where σ is the standard deviation of the impact distribution and n is the number of observations.

period. The profile plot [6] depicting market impact over participation rate suggests a slightly concave curve, as it is also taken by Almgren *et al.* (2005). As mentioned earlier we provide two alternatives to explain this relation between participation rate and market impact. The first proposal is a simple linear model and the second is a power law model.

3.1 Linear Model

Due to the technical reasons, a linear model may be preferable for some applications (see for example chapter 4). Additionally, it has the advantage, that it can easily be implemented in a portfolio optimization without increasing the complexity of the problem (see appendix A). It also can be motivated by the fact that it is the first term of a Taylor expansion and a good approximation for a small range of the participation rate. The linear model is given by the following function

$$M(v) = m * v + b \tag{6}$$

where M is the market impact and v the participation rate with the parameters m and b.
We use the maximum likelihood method for the parameter estimation and an asymmetric Laplace-distribution for the residuals r.

$$r(x) = \frac{\tau(1-\tau)}{\sigma} \cdot e^{-\rho_\tau(\frac{x-\mu}{\sigma})} \tag{7}$$

where ρ_τ is given by

$$\rho_\tau(u) = \frac{|u| + (2\tau - 1)u}{2} \tag{8}$$

The parameters of the linear model (equation 6) m and b are estimated separately for every market (Europe, Japan, USA, Canada) because it is reasonable to assume that the different market characteristics lead to different market models. Indeed this effect can be observed in the fit results (see table 2 and figure 4).

[6] average impact per bin of participation rate, see http://root.cern.ch/root/html/TProfile.html

Table 2: Fit results of the linear model

	scale, m	intercept, b
Europe	55.924 ± 8.837	0.938 ± 0.197
Japan	123.377 ± 9.810	2.235 ± 0.394
USA	77.827 ± 6.149	1.797 ± 0.129
Canada	114.02 ± 22.890	0.997 ± 0.280

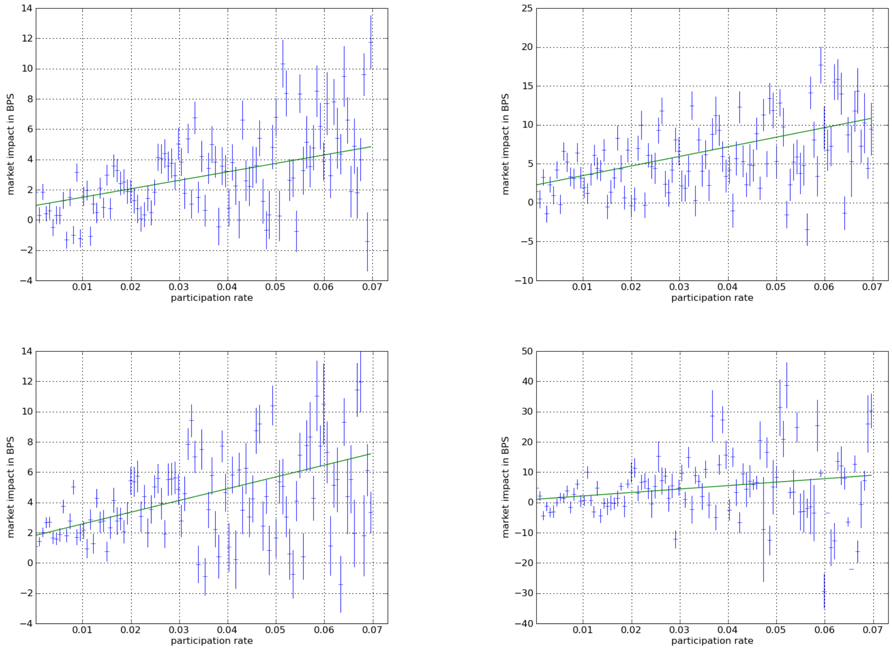

Figure 4: linear market impact model, market impact over participation rate; top left to bottom right: EU, JP, USA, CN

3.2 Power Law Model

The various profile plots market impact over participation rate suggest a slightly concave model, which is also in line with existing literature (see for example Almgren *et al.* (2005)). Especially for larger ranges of the participation rate, the concave model fits the observations much better. Therefore we propose a power law model defined by:

$$M(x) = m * v^a + b \tag{9}$$

Here, we suggest two slightly different interpretations of the power law model describing the market impact for the different markets. The first is similar to the one of the linear model, where all parameters are estimated separately for each market. The results of this fit procedure can be found in table 3.

Table 3: Fit results of the power law model each market fitted separately

	scale, m	exponent, a	intercept, b
Europe	12.992 ± 8.362	0.437 ± 0.222	-0.247 ± 0.947
Japan	40.275 ± 24.602	0.511 ± 0.208	-1.112 ± 1.427
USA	30.320 ± 21.972	0.686 ± 0.210	1.231 ± 0.374
Canada	20.061 ± 15.685	0.423 ± 0.241	-0.221 ± 1.327

Our second approach is different in the sense that we estimate parameter m and b separately for each market but the exponent a is estimated together for all markets. This has the advantage that the complete data sample can be used to estimate the parameter, leading to lower statistical uncertainty of the fit results. The likelihood function of this fit can be written as

$$L = \prod_i L_i \left(\vec{v}_i, \vec{y}_i, \vec{p}_i\right) \tag{10}$$

where i represents the different markets (EU,US,JP,CN). The parameter set of market i is given by $\vec{p}_i = (m_i, a, b_i)$. L_i denotes the likelihood function for one market which can be written as

$$L_i \left(\vec{v}_i, \vec{y}_i, \vec{p}_i\right) = \prod_{j=0}^{N} r(M(v_i^j) - y_i^j) \tag{11}$$

where $\vec{v}_i$ is the sample of participation rates and $\vec{y}_i$ are the price changes in market i. $M(v)$ is defined by equation 9 and $r(x)$ is the residual distribution, see equation 7.

Furthermore, it seems reasonable to take the same exponent for more than one market because (see table 3) the estimated exponents are quite similar in all 4 markets. They do not differ significantly on a 95 % significance level, when using the error propagation on the difference of the estimated parameters with their uncertainty [7].

The results of the combined fit including all 4 markets (EU, US, JP, CN) can be found in table 4. The exponent is, by definition, the same for all 4 markets and is estimated as 0.534 ± 0.115.
It is remarkable that there is evidence for a negative intercept of the Japanese market. A negative intercept in this model is not reasonable because this means that very small trades create negative costs which would imply the possibility of arbitrage. This fact and the a priori knowledge that the Japanese market behaves different than the other markets with respect to regulations (such as up-tick rule (see 2.3)) motivates us to do the fit again without the Japanese market. The effect of the negative intercept can be explained by the up-tick rule because short-sells have an execution probability significantly lower than 1. So the up-tick rule which affects only the short-sells in the Japanese market leads to a bias which can be explained as follows: If the execution price is higher than the arrival price, a very high percentage of the short sells should be executed and the measured market impact for these trades is negative. If the execution price is lower than the arrival price, the execution probability is worse (fewer shares are traded) and the measured market impact is large. So there exists a bias towards lower market impact in the Japanese market using the current method to estimate market impact. In spite of the knowledge about this effect, we are not able to remove it and integrate the Japanese data in the analysis. Assuming removal of the short sells for the Japanese market, this bias would be lost, but another bias may appear: the assumptions for the negligence the market movement in section 2.2 would be hurt.
The results of the fit with the European, US, and Canadian market can be found in table 5. Figure 5 shows the corresponding plots.

[7] The difference of the estimated parameters for different measurements is calculated. The error of the difference can be estimated with the help of error propagation, see equation 12, 13 and 14.

$$\mu = \mu_1 - \mu_2 \tag{12}$$

$$\sigma^2 = \left(\frac{d\mu}{d\mu_1} \cdot \sigma_1\right)^2 + \left(\frac{d\mu}{d\mu_2} \cdot \sigma_2\right)^2 \tag{13}$$

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 \tag{14}$$

Table 4: Fit results of the power law model Europe, Japan, USA and Canada together

	scale, m	exponent, a	intercept, b
Europe	15.872 ± 6.017	0.534 ± 0.115	0.095 ± 0.521
Japan	42.939 ± 15.577	0.534 ± 0.115	-0.980 ± 0.680
USA	18.996 ± 7.160	0.534 ± 0.115	0.957 ± 0.344
Canada	26.816 ± 11.432	0.534 ± 0.115	0.256 ± 0.807

Table 5: Fit results of the power law model Europe, USA and Canada together

	scale, m	exponent, a	intercept, b
Europe	16.345 ± 7.500	0.547 ± 0.143	0.131 ± 0.557
USA	19.801 ± 9.412	0.547 ± 0.143	0.982 ± 0.389
Canada	27.774 ± 13.726	0.547 ± 0.143	0.302 ± 0.860

Comparing the statistical uncertainties of the parameter estimations of table 2 on the one hand and tables 3, 4, and 5 on the other hand, it is conspicuous that the scale parameters of the first table are estimated much more significantly than in the remaining tables. This is explained by the error estimation method of the parameters and their correlations. The model parameters scale, intercept, and exponent are correlated. When slightly varying one of the model parameters, a solution for the remaining parameters can be found describing the data set almost as good as the optimal solution. The errors of parameters estimated by the maximum likelihood procedure can be estimated by varying one parameter until the likelihood function rises by 0.5. During the variation of this parameter, for all other fit parameters the maximum of the likelihood function has to be found (see Blobel and Lohamann (1998), p. 189-191). We have done a thorough analysis of the regression errors and being rather conservative in the error estimate. To convince the reader of the significance of the findings we test the two Null-Hypotheses: keeping on one hand the intercept equal 0 the slope is significantly different from 0. And likewise on the other hand if the slope is kept equal to 0 the intercept is significantly different from 0.

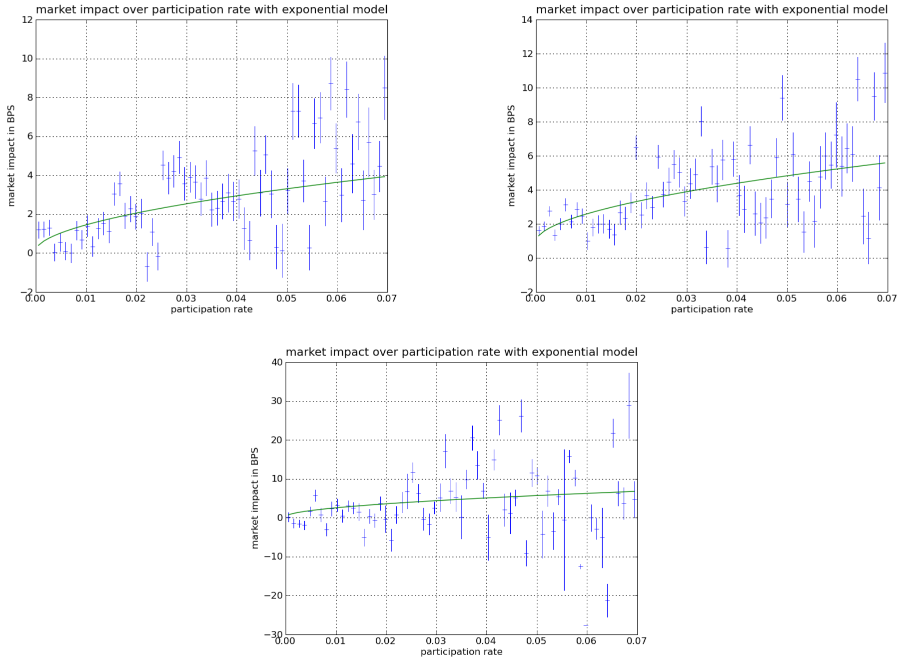

Figure 5: power law market impact model, market impact over participation rate; top left to bottom right: EU, USA, CN

4 Individualized linear regression analysis

In this section we try to improve the simple linear regression model using an algorithm described in Scherrer *et al.* (2010). The algorithm is an advanced linear regression model in which the slope and the intercept of the regression are allowed to be dependent on every single event. This model applied to the impact analysis generalizes the equation of the simple linear regression

$$t_i = mv_i + b + \epsilon_i \tag{15}$$

where ϵ_i is the residual of event i and v_i is the volume fraction, to

$$t_i = m(x_{1,i}, ..., x_{n,i})v_i + b(x_{1,i}, ..., x_{n,i}) + \tilde{\epsilon}_i \tag{16}$$

The slope m and the intersect b are not constant any more and can depend in a nonlinear way on the variables $x_1, ..., x_n$.
This ansatz is reasonable because already the simple linear regression describes the impact quite well but we would like to understand the corrections to the linear model with respect to some external variables such as volatility of the specific stock, tick sizes, market capitalizations etc.
We use the individualized linear regression instead of the power law for some practical reasons. A market impact model may be used in a portfolio optimization performed by a trader. The trader takes the predictions for the stock market returns into account, but his trades in turn will influence the stock returns. A portfolio optimization has to be fast and the solver for the optimization problem is much faster for a linear impact model.

In the next two sections we will describe the basic idea of the algorithm and present the results and improvements compared to the simple linear regression model.

4.1 The individualized linear regression algorithm

The details of the individualized regression analysis are described in Scherrer *et al.* (2010).
The first step is to transform the input variables $x_1, ..., x_n$ to be uniformly distributed. This means, by definition, that if we plot a histogram of the specific input variable we have the same amount of statistic in every bin. In the next step we divide the input variable into k bins. The parameters m and b and their errors are estimated for every bin of the input variable. In order to make the algorithm robust against statistical fluctuations we use additionally a spline fit to smoothen the dependencies of m and b on the specific input variable.
This procedure is done for all input variables.

For one event i we get one prediction for m_i and b_i for each input variable. That means we have n predictions for m_i (b_i). We would like to end up with **one** prediction only for m_i (b_i) of a certain specific event. The easiest ansatz would be to average the m's and the b's to get

$$m_i = \frac{1}{n}\sum_{j=1}^{n} m_j \qquad b_i = \frac{1}{n}\sum_{j=1}^{n} b_j \tag{17}$$

But this choice is not optimal. The prediction coming from a variable with a high correlation to the target t_i should have a larger weight than the prediction coming from a weakly correlated variable.

A problem could also appear if vector $\vec{x}$ is introduced in which all the components are highly correlated to each other. The algorithm should recognize such correlations and make sure that the statistical significance of the correlation between the input variables and the target is not increased by introducing further redundant variables which are highly correlated to the rest of the variables.

Obviously we would like to use an algorithm which can deal with correlations among the input variables and which is able to decide if a variable has a statistically significant correlation to m (b) at all. If there is a large correlation of a variable and m (b), the weight of the estimator should be larger than the weight given to an unimportant variable. And if the input variables are correlated among each other the algorithm should treat these correlations correctly.

For this kind of problem we can use the NeuroBayes®software [8] which is described in Feindt (2004). We use the n predictions for m_i, the n predictions for b_i and the variable v_i as input vector (details see Scherrer *et al.* (2010)). The target is defined by the execution price.

4.2 Input variables of the individualized linear regression

As we do not have enough statistics for the Canadian market, we only take the Japanese, the European and the US market into account.
To understand the underlying dynamics of execution price and impact we introduce some appropriate variables (see table 6) and analyze the improvements to the simple linear regression model. This analysis should describe all dependencies of the pa-

[8] Developed by Phi-T® Physics Information Technologies GmbH

Table 6: Input variables for the individualized linear regression analysis

input variables
market-return-arrival-close
dir-market-return-arrival-close
dir-market-return-eq
vola
rel-ts
liquidity
t
volume-fraction
market-cap

rameters m and b on the input variables. Therefore we will use not only variables which only include past information, but also variables which include information on the future.
The execution price for a specific stock is firstly dependent on the return of the stock which would have taken place without the order of the market participant and secondly on the impact of the order, according to equation (3).
Note that the return of the stock which was traded must not be used as input variable, because the return already includes its impact. Therefore it is not reasonable to explain the impact using an input variable which already includes the impact. However, we can neglect the impact of the order on the underlying stock market index (EuroStoxx 50 for Europe, S&P 500 for the USA, Nikkei for Japan) and instead define an input variable using the index return.
Let us define the following variables for the model:

- "market-return-arrival-close" denotes the return of the stock market index in the trading period
- "dir-market-return-arrival-close" denotes the return of the stock market index multiplied with the direction the stock was traded (+1 for a buy and −1 for a sell)
- "dir-market-return-eq" is defined by the stock market index return in the trading

period multiplied with an estimator of the beta-factor[9]. The beta-factor is estimated from historical data and does not include information of the future.

- "vola" is an estimator for the volatility
- "rel-ts" is the relative tick size of the stock (tick size divided by the arrival price when the trading period begins)
- "liquidity" is defined by the traded volume in Euro (for the specific stock) at the trading day
- We introduce the variable "t" reflecting the date to account for an explicit time dependence of the parameters m and b.
- The variable "volume-fraction" includes the relative traded Euro-volume of the specific stock compared to all traded stocks at that day for the market in which the stock is traded (Europe, Japan, USA).
- Finally for the market capitalization we define the variable "market-cap".

4.3 Results of the individualized linear regression

Our findings suggest that the variables "market-return-arrival-close", "liquidity", "volume-fraction", "market-cap" and "t" do not have any significant correlations to the parameters m and b.

Due to the fact that the relative frequencies of buy and sell orders are equal, it is not surprising that this variable is not important. Much more important is the return weighted with the trading direction ("dir-market-return-arrival-close"). The basic idea why this variable has been introduced is that the return of the asset is generally correlated to the market portfolio. The influence of the market participant on the market portfolio is negligible, so this variable should be a good estimator for the return of the asset in the trading period as it contains information on the future. We include it in the analysis to understand the underlying market components which influence the execution price, but it cannot be used for prediction.

[9]The factor β is defined for an asset i in the CAPM (see Sharpe (1964)) as $\beta_i = \frac{cov(r_i, r_M)}{\sigma^2(r_M)}$ where r_i is the return of the asset i and r_M is the return of a market portfolio.

All information of the variables "liquidity", "volume-fraction" and "market-cap" is completely absorbed in the participation rate.

The most important effect in all markets investigated is a high correlation of the parameters m and b to both the index-return weighted with the trading direction (see figure 6) and to the index-return weighted with the trading direction and the beta-factor (see figure 7). The dependence of the b-parameter on these two parameters is much more significant than the dependence of the m-parameter.

In the algorithm the last bin of the plots has a special meaning (see Scherrer *et al.* (2010)): If some input variables are not known or believed to be wrong for some events, one can activate a special flag for these events. Hence there are two different possibilities:
The first possibility is to define a variable that is not known (or not correct) for all events. Consequently those events which are not known (or incorrect) are separated in the last bin. The parameters m and b are then estimated for this input variable bin and afterwards not included in the spline fit [10]. The second possibility is that there are no events in the training sample which are filled in the last bin. Then the estimator of m (b) is defined by the mean of all other bins.
If the user has adjusted the parameters of the prediction model on historical data and would like to use the results to forecast an event in which the variable is not known (or wrong), the estimator of the special bin is used.

In figure 8 we see that the parameter b is significantly correlated to the relative tick size but only in the US and in the Japanese market. A possible reason could be that the definition of the tick size in Europe is relative to the price level (see e.g. the tick size structure at XETRA in table 7) whereas in the USA tick size is absolute 1 cent and constant for all stocks. The tick size definitions in the Japanese market lead to extremely large relative tick sizes (tick size over stock price) for some stocks. This is also valid for US stocks with a low absolute stock price. It is different to the European market where the tick size depends on the price of the stock (the rules are similar for all European exchanges). This leads to quite small relative tick sizes for all European stocks. Maybe the European tick size definition is responsible for the

[10] It has a completely different meaning compared to the rest of the bins, so the assumption that m and b are smoothly depending on the input variable is not valid for the last bin.

Table 7: Structure of the tick sizes in Europe.

price of share	tick size
0 EUR - 9.999 EUR	0.001 EUR
10 EUR - 49.995 EUR	0.005 EUR
50 EUR - 99.99 EUR	0.01 EUR
100 EUR - ∞	0.05 EUR

independence of tick size and execution price.

The parameters of the model are also slightly dependent on the volatility but only in Japan and in the USA (figure 9). While the parameter m is fairly constant and independent of the volatility, the b parameter is correlated to the volatility.

As mentioned earlier the goal of the individualized linear regression applied in this chapter is to find the importance of the underlying factors which are responsible for the impact. We compare the mean absolute deviation (mad) and the standard deviation σ which are defined as follows:

$$mad = \frac{1}{N}\sum_{i=1}^{N} |m(x_{1,i}, ..., x_{n,i})v_i + b(x_{1,i}, ..., x_{n,i}) - r_{s,i}| \quad (18)$$

$$\sigma = \frac{1}{N}\sum_{i=1}^{N} (m(x_{1,i}, ..., x_{n,i})v_i + b(x_{1,i}, ..., x_{n,i}) - r_{s,i})^2 \quad (19)$$

In table 8 we summarize the results of the simple linear regression and the individualized linear regression. The last column is the relative improvement of mad and σ if the individualized linear regression is used.

For the European market we can improve the simple linear model approximately by 11%, while the US market and the Japanese market are improved by 8.5% and 6.5% respectively. In our analysis we found that this effect is mainly based on the return of the underlying stock market index as long as there is a dependence of the relative tick size and the volatility in the US and the Japanese market.
If a trader would like to use a market impact model for the portfolio optimization, he would have to estimate the volatility and the volume which will be traded during

Table 8: Comparison of the results coming from the simple linear, the individualized linear regression model, and the model based on power law.

	simple linear regression	individualized linear regression	relative improvement of ind. lin. reg compared to simple lin. reg	power law (separate markets)	power law, markets fitted together
mad_{EU}	25.114	22.168	11.7%	25.116	25.115
σ_{EU}	34.608	31.053	10.3%	34.613	34.612
mad_{US}	16.307	14.726	9.7%	16.308	16.307
σ_{US}	23.657	21.910	7.4%	23.655	23.655
mad_{JP}	22.634	21.014	6.2%	22.534	xxx
σ_{JP}	30.529	28.513	6.6%	30.583	xxx
mad_{CN}	22.813	xxx	xxx	22.035	22.034
σ_{CN}	31.510	xxx	xxx	30.579	31.575

the trading period. These input parameters can be estimated quite accurately from historical data. The relative tick size is also known before the trading period.

We have thus explained that the beta weighted return of the underlying stock market index during the trading period is extremely important for the impact. This variable decreases the variance of the residuals quite dramatically. But it is a problem to estimate the return of the stock market index during the trading period a priori. A trader can either have a mathematical model for the index return and utilize this for the impact model. Alternatively he could relinquish the variable at all, which would lead to a larger variance of the distribution of the residuals.

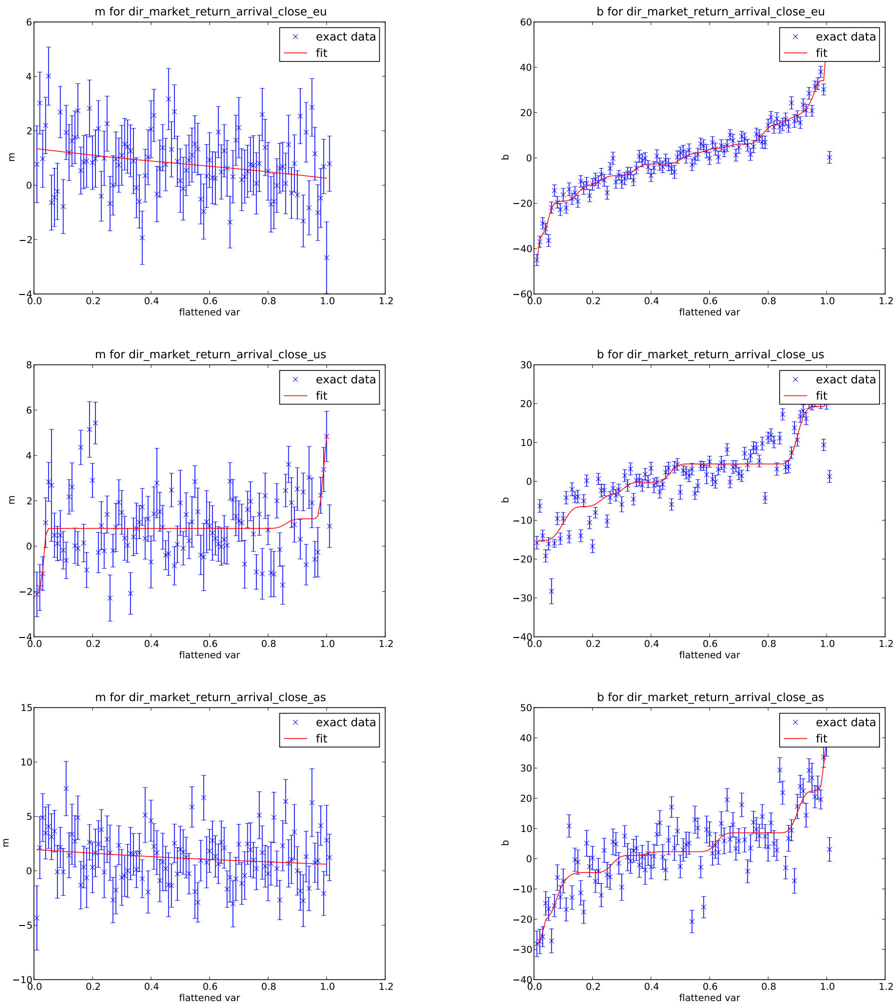

Figure 6: Correlations of the parameters m and b to the stock market return of the underlying index (see text) weighted with the direction of the trade.

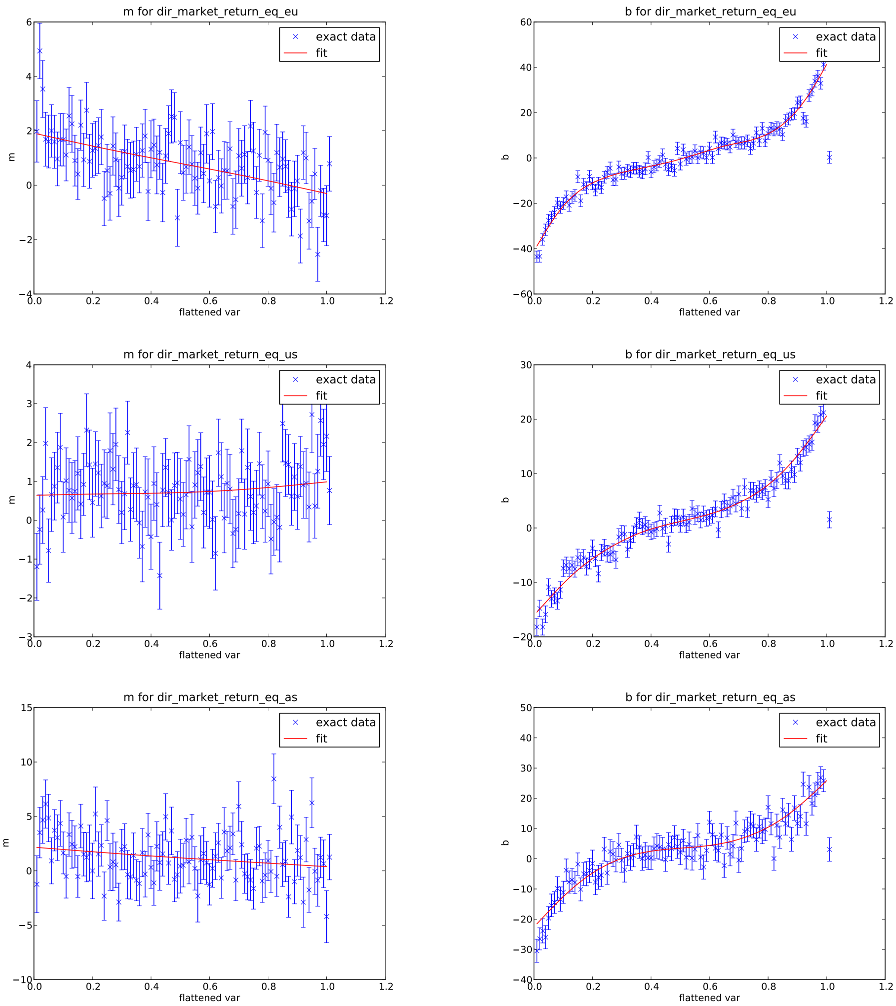

Figure 7: Correlations of the parameters m and b to the stock market return of the underlying index (see text) weighted with the direction of the trade and the beta-factor.

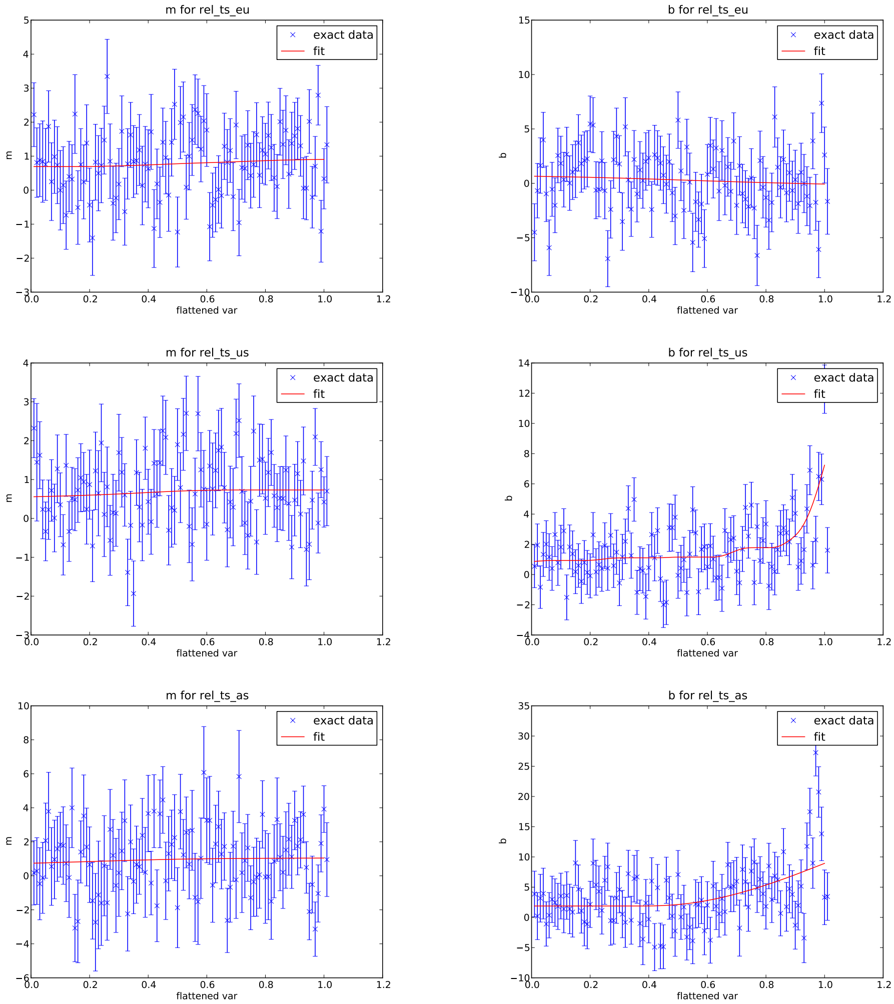

Figure 8: Correlations of the parameters m and b to the relative tick size of the stock.

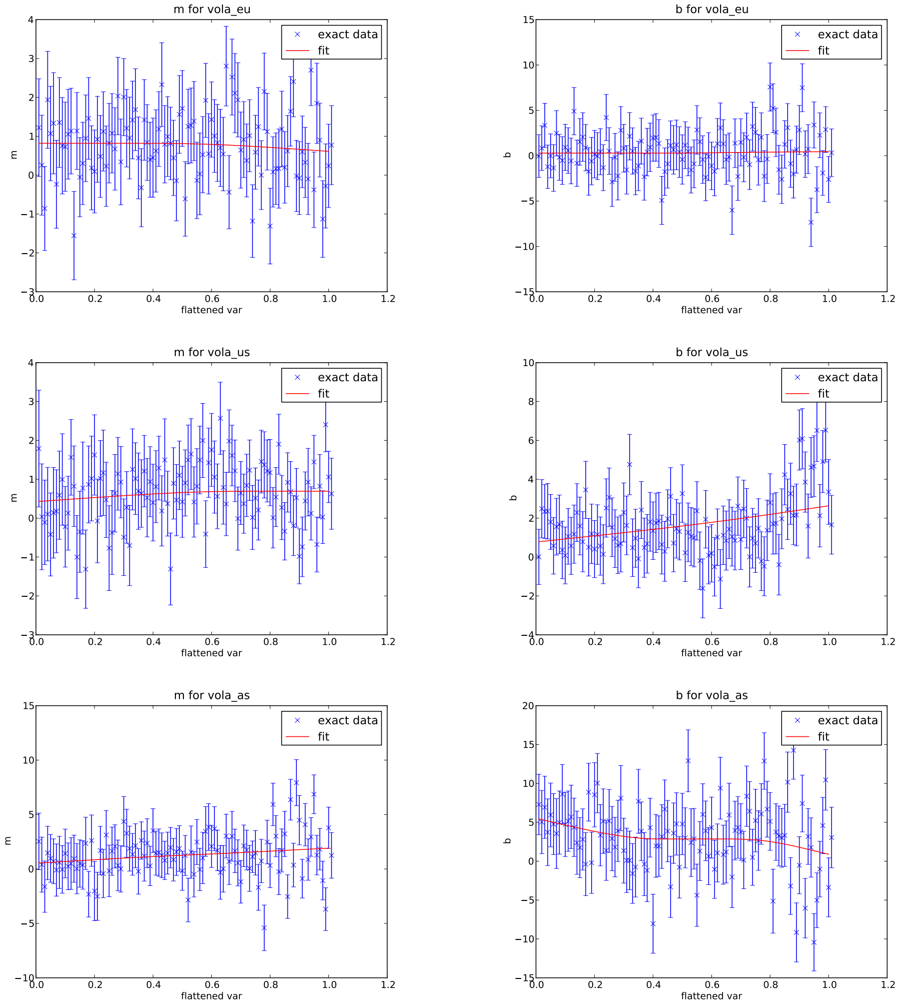

Figure 9: Correlations of the parameters m and b to the volatility of the traded stock.

5 VWAP - The Optimal Trading Strategy

With the knowledge of the market impact models and the dependence of the market impact on the participation rate, it is obvious to have a look on the optimization of the trading strategy. We assume that the functional form between participation rate and market impact is the same for different trading periods. For our three types of models (the linear, the power law, and the individualized linear model) we can show that the VWAP strategy is the optimal execution strategy.

We demonstrate this by splitting the trading period in N sub-periods and varying the volume executed in each of the sub-periods. The market impact is optimal with the constraint of full execution inside the given time period. This may be written as follows:

$$f = \sum_{i}^{N} \left(m \cdot \left(\frac{v_i}{V_i} \right)^{\beta} \cdot v_i + b \cdot v_i \right) + \lambda \cdot (\sum_{i}^{N} v_i - v) \tag{20}$$

V_i is the entire traded volume of the current stock in time period i, v_i is the volume traded by the special algo in i and m, β, as well as b are the model parameters. Parameter λ is the Lagrange multiplier of the constraint of full execution and v the size of the algo order which has to be executed.

The result of the optimization of the entire market impact is:

$$v_i = \frac{V_i}{V} \cdot v \qquad \forall i \in [1, 2, ..., N] \quad and \quad \beta \neq 0 \tag{21}$$

and

$$\frac{v_i}{v_j} = \frac{V_i}{V_j} \tag{22}$$

This can be interpreted in the way that the volume of the algo order should distributed over the given period of time proportional to the entire trading volume of the stock, which is exactly the idea of a VWAP trading algorithm.

6 Conclusion

The current work investigates the variables describing the market impact of orders at stock markets executed by a trading algorithm. We find the by far most important variable to be the participation rate. Order executions in four markets (Canada, Europe, Japan, and USA) are examined. The Japanese market behaves different with respect to the other markets due to the up-tick rule for short-sells. The market impact is biased towards lower values because the execution probability of short-sells is small in bear markets if the up-tick rule is valid.

This work provides a pragmatical approach of measurement which is rarely done in the literature. The reason for the latter may be that most publicly available data sets cannot be utilized because any observed order cannot be related to the market participant and its consequential impact. Only if the set of (small) sub-orders belonging to a (large) algo order can be identified as a whole and followed up, the market impact of large transactions can be measured.

A linear model describing the dependence between the participation rate and the market impact is provided for all four markets. The linear model as input for a portfolio optimization has the advantage that the optimization function is quadratic in order size which is much faster to solve. In order to improve the linear model on the one hand and keep it linear in order size on the other hand, an individualized linear regression algorithm is introduced. This algorithm allows to handle additional description variables in the model. The stock market index movement during the trading period of the algo order is a quite important variable to describe the market impact. It is not possible to predict the market impact with the help of the market movement because this would contain future information, but it helps to better understand the dynamics. The market movement of the trading period is of course not known when trading starts. The volatility of the stock price is also slightly correlated to the market impact. For the US and the Japanese markets we see a tick size dependency. This is not true for the European market. Thus the tick size definitions of the European markets seems to be more efficient in that sense.

The comparison of the different models shows that there is no large difference in the reduction of the mad and σ of the residuals except the individualized linear regression

due to the stock market movement as a input variable. The linear models works well for a narrow range of the participation rate. Having a wide range of participation rates, it gets obvious that the functional dependence between participation rate and market impact is not linear anymore but concave. Therefore we propose a power law as it is already done in literature. For the combined measurement (EU,USA,Canada), we estimate the value of the exponent of the model to be 0.547 ± 0.143. The result is in line with other similar measurements. It can be utilized to verify the considerations of the market microstructure theory.
Additionally we have shown that the VWAP trading strategy is the optimal execution strategy for all the discussed market impact models.

A Portfolio Optimization

The general objective of a portfolio optimization is the maximization of the expected profit while keeping risk constant. Having predictions and the estimated risk, an optimization software has to find the best allocation. The actual asset allocation before the optimization is started is denoted by $\vec{a^{(0)}}$, whereas $\vec{a}$ is the optimal asset allocation to be determined. The expected earnings are given by

$$f(\vec{a}) = \sum_{i=1}^{n} \left(a_i \mu_i - I(a_i - a_i^{(0)}) \right) \tag{23}$$

where μ_i is the expected return of asset i and I is the impact depending on the traded volume of the asset.
The market impact model describes the relative impact of an order. The calculation of the expected profit takes into account the absolute costs of the impact. Alltogether the relative impact, coming from the impact model, is multiplied by the traded volume leading to:

$$f(\vec{a}) = \sum_{i=1}^{n} \left(a_i \mu_i - (a_i - a_i^{(0)})(m(a_i - a_i^{(0)}) + b) \right) \tag{24}$$

Using the linear impact model, the optimization problem itself remains quadratic (QP). Using the power law, we end up with a general nonlinear problem. Optimization algorithms solve a linear (LP) or a quadratic problem much faster than general nonlinear problems. In a high frequency trading set-up, the optimization algorithm has to be fast. Therefore the linear impact model has the great advantage of not increasing the complexity of the optimization problem.

References

Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct estimation of equity market impact. *Risk magazine*.

Blobel, V. and Lohamann, E. (1998). *Statistische und numerische Methoden der Datenanalyse*. Teubner Studienbuecher.

Demsetz, H. (1968). The cost of transacting. *The Quarterly Journal of Economics*, *82*(1), 33–53.

Feindt, M. (2004). A neural bayesian estimator for conditional probability densities.

Fränkle, J. and Rachev, S. T. (2009). Review: Algorithmic trading. (1), 7–20.

Perold, A. (1988). The implementation shortfall: Paper versus reality. *Journal of Portfolio Management*, *14*, 4–9.

Scherrer, C., Rachev, S. T., Feindt, M., and Fabozzi, F. (2010). From the simple linear regression to the individualized linear regression.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, *19*, 425–442+.