Essays on Prediction and Behavior of Financial Markets

Kumulative Inaugural-Dissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

an der Wirtschaftswissenschaftlichen Fakultät
der Universität Augsburg

vorgelegt von

Martin Niemann

2017

Erstgutachter:	Prof. Dr. Dr. h.c. Günter Bamberg
Zweitgutachter:	Prof. Dr. Yarema Okhrin
Vorsitzender der mündlichen Prüfung:	Prof. Dr. Hans Ulrich Buhl
Datum der mündlichen Prüfung:	12. Januar 2017

Table of Contents

Index of Essays

This dissertation presents the following published or under review essays:

E.1 Niemann, M., Neukirchen, M., Schmidt, J. H. 2008. *Improving performance of corporate rating prediction models by reducing financial ratio heterogeneity.* Published in: Journal of Banking & Finance 32, 434-446, DOI: 10.1016/j.jbankfin.2007.05.015 (primary author, 50% share of authorship – VHB-JOURQUAL 3: A, VHB-JOURQUAL 2.1: A)

E.2 Niemann, M. 2015. *Exploiting attention-driven mispricing: Evidence from actual-dollar trading.* Presented on Jan 29, 2015 at the Institute of Statistics, University of Augsburg, to Prof. Dr. Yarema Okhrin and Prof. Dr. Dr. h.c. Bamberg.

E.3 Niemann, M. 2015. *High Frequency Trading Intensifies Intraday Extreme Events in Stock Returns.* Currently in resubmit in: Journal of Business Economics (formerly ZfB Zeitschrift für Betriebswirtschaft – VHB-JOURQUAL 3: B, VHB-JOURQUAL 2.1: B). Presented on Jan 29, 2015 at the Institute of Statistics, University of Augsburg, to Prof. Dr. Yarema Okhrin and Prof. Dr. Dr. h.c. Bamberg.

I. Introduction

This cumulative dissertation develops and applies methods to predict and empirically study financial market behavior. It presents three papers examining different research questions on the economic and statistical laws governing financial markets. Despite the variation in contexts, a common theme is developing robust prediction models for the questions at hand. The underlying motivation is to support sound decision-making, by identifying the most important forces at work in a given system, extracting the most powerful predictive features from the data, and combining them in a well-specified algorithm. In sum, this work is as much on building robust predictive models as it is on solving the specific research questions on financial market behavior.

Each of the three studies has been motivated by questions arising in different contexts of my business career. Working part-time on the dissertation for over nine years provided opportunities to approach financial markets from different angles: First as a management consultant focusing on risk management in financial institutions, second as co-founder of a quantitative asset management firm, and third as an entrepreneur developing data analytics solutions for decision-making.

The first study, *Improving Performance of Corporate Rating Prediction Models by Reducing Financial Ratio Heterogeneity*, develops a methodology to construct better performing models to predict credit default rates of large corporations across different industries. It was motivated by the fact that our consulting team had difficulties to construct rating models for large corporates, due to limited available data on defaults and heterogeneity in financial ratios across industry groups. Published work did not provide much methodological help. This motivated developing our own methodology to account for industry heterogeneity within the rating model, and thereby achieving a notable improvement in prediction accuracy.

The second paper, *Exploiting Attention-driven Mispricing: Evidence from Actual-Dollar Trading*, develops a systematic trading strategy for U.S. stocks and successfully trades it in a true out-of-sample test with real money. These results not only motivated investors to provide the seed funding to start a quantitative asset management firm, it also posed the question of how these profits could be possible and persistent for a longer period. Given that the widely accepted efficient market hypothesis (Fama, 1970) implies that financial markets eliminate such profit opportunities quickly, this conflicting observation deserved further investigation.

Third and finally, the essay *High Frequency Trading Intensifies Intraday Extreme Events in Stock Returns* investigates whether high frequency trading (HFT) activity exacerbates large intraday price moves in the stock market. The idea of investigating the link between HFT and intraday extreme events was motivated by my intraday market observations from countless hours of automated trading surveillance. Thereby, sudden bursts of activity and volatility – often without any news – were a surprisingly regular phenomenon. At the same time, there is a dichotomy in the literature. On the one hand, several published empirical studies indicate that HFT activity dampens volatility and improves market quality. On the other hand, theoretical models and institutional traders formulate multiple plausible mechanisms by which HFT could cause extreme events in short-term stock returns.

Overall, although the contexts and topics have changed considerably over time, all papers share an underlying research motivation: obtain a better understanding of the economic and statistical laws governing the behavior of financial markets, and to do so, build "good models" that derive these laws from empirical data. Ultimately, the aim of all analysis is to make sound decisions. Across the three studies, working towards this goal has led to shared methodology and techniques. The next sections of this introduction will present these commonalities and summarize the three essays in the context of these themes and their contributions to their respective bodies of literature. The introduction concludes with overarching learnings on financial market behavior from the three papers. The core of this dissertation in sections II to IV contains the actual essays. Section V discusses their contributions in light of new publications in their fields and derives implications for further research.

1. Building Robust Predictive Models

A good model has statistically significant predictive power out-of-sample. Not only does it fit well the empirical data it has been developed and trained on, but it also stays robust when making predictions on new data. This determines whether a model can support real-world decisions. This section presents a result-focused framework which subsumes the most important requirements defined in prior literature.[1] Thereby, *achieving predictive power* and *achieving robustness* are the main goals.

[1] Academic literature and published comments by researchers offer several related frameworks of the characteristics and requirements of best practice empirical models. For instance, Chipperfield (2013) proposes the criteria fit, predictivity, parsimony, and sanity. While the first two essentially subsume the outcomes of predictive power and robustness on new data, the latter two represent key requirements to achieve these outcomes. Parsimony calls for selecting the simplest well-performing model, which supports robust out-of-sample performance as well as traceability of the model variables and parameters. Sanity requires a well-specified model in terms of using a link function or classification algorithm which fits the data and the phenomenon to be modeled. Furthermore, a "sane" model ensures that all assumptions by the chosen methodology be met. In addition, researchers should apply common sense to exclude nonsensical results. In an extensive survey, Kuntz et al. (2013) extract requirements for decision models in a large survey of literature as well as expert panels. They categorize requirements into structure, data, consistency/validation, and communication. Structure requires to pick the right classification algorithm for the problem, and to assure that all assumptions are met. Data includes the choice of appropriate modeling samples, specification of parameters as well as factor transformation (e.g. winsorizing variables). Consistency/validation calls for building on established learnings and principles from previous studies, as well as predictive validity of the model out-of-sample. Communication requires traceability and parsimony of the model, as well as transparency through thorough description and documentation of a model.

Both goals pose several key requirements. These criteria represent by no means a comprehensive checklist of how to build a good model; rather, the aim is to illustrate the most important drivers to achieve robust predictive performance.

1.1. Predictive Power

Achieving predictive power requires specifying the right model *structure and classification algorithm*, and *adapting it to the task environment* by extracting significant predictors from the data and using or optimizing statistical circumstances and properties to increase statistical power.

Correctly-specified structure / prediction algorithm – choosing the right *prediction algorithm* or *link function* is a key driver of model performance. The first task is not to make an artisanal error in model specification. For example, if the dependent variable is bounded by zero, simple linear regression might not be the right algorithm, as it could predict negative values. Instead, a logistic link function might be appropriate (see e.g. Chipperfield, 2013). Furthermore, the algorithm should make reasonable assumptions about the data. For instance, a common pitfall in time-series regression is that significances of parameters are biased upwards if errors are auto-correlated. This problem can be solved by adjusting the regression specification (e.g. by adding lagged dependent variables) or might require more advanced model types. The second task is choosing an algorithm with superior capability to learn the laws of the system we study. If relationships between our independent predictors and our variable of interest are significantly different from linear, ordinary linear regression will at best return imprecise estimates. Linearizing predictor variables (e.g. through log-transformation) might fix this, or we might have to switch to a different algorithm altogether. In conclusion, the assumptions implied by our choice of model type should fit to the data and the system we study (or at least not be too strongly violated), and the algorithm has to be able to capture the underlying relationships.

Adapted to task environment – Todd & Gigerenzer (2000), who study the performance of decision heuristics[2] in different environments, define a well-adapted model as *ecologically rational*, in that it fits its use of information and logic to the information structure in the environment. First of all, this means using underlying laws of information in the system we study to extract powerful predictors. To illustrate, assume we have to solve the question which of two randomly picked cities is the larger one. An ecologically rational predictor variable would be to ask whether we know the cities. If we only know one of the two, this indicates the larger city with a high accuracy. This concept is powerful because mere recognition of a city is highly correlated to the many cues that predict city size: have we heard from it in the news, does the city have a renowned university? All these criteria are correlated to city size and are partly subsumed by the simple predictor. It is powerful because it fits the structure of the socio-economic system it tries to predict. Next to the definition of good predictive factors from available data, a good model uses the statistical circumstances to its advantage. For instance, Fama & Macbeth (1973), in their seminal paper on estimating equity risk premia with a two-step regression, make use of the fact that estimates of second moments (variances and co-variances) need much less data than the first moment (the mean) to converge towards small confidence intervals (Chacko et al., 2014). The procedure is still used by researchers today (see e.g. Barinov, 2015; Conrad et al., 2015; Khovansky & Zhylyevskyy, 2013). In sum, *adaptation to the task environment* can be done in two ways: *fundamentally*, we can

[2] Todd and Gigerenzer focus on the performance of heuristics; however, since every heuristic is a simple model of the system it is applied in, the implications apply for our discussion of models as well.

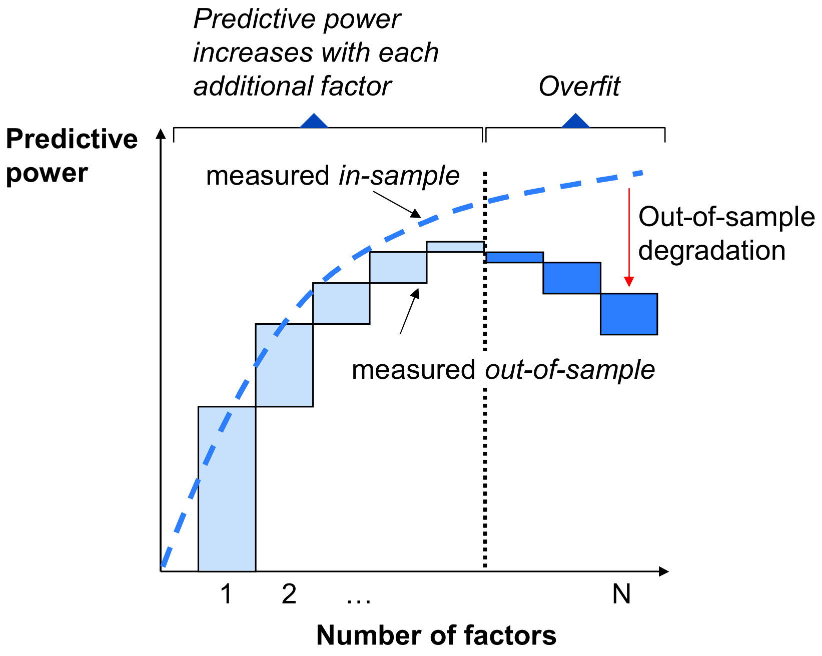
define powerful predictive factors which exploit the information structure of the system we study; *statistically*, we can utilize the properties of the underlying data and mathematical laws to our advantage.

1.2. Achieving Robustness

A robust model performs well out-of-sample. Facing new data, it makes right predictions on average, because it has generalized the most relevant laws of its task environment. There are two drivers of out-of-sample robustness: first, ensuring *parsimony and simplicity* during model construction, and second, performing *validity and robustness checks* to evaluate candidate and final models.

Parsimony and simplicity – when we face the choice between two models with approximately the same explanatory power, we should always choose the simpler one (Chipperfield, 2013). However, this is easier said than done: a researcher fitting an algorithm to a data sample achieves subsequently better measures of predictive power for each additional variable included in the model. Eventually, this results in *overfit*. Then, a model's measured predictive power on the estimation sample overstates what it can realistically achieve facing new data. Each new *free parameter* allows describing the underlying laws of the studied environment in more detail; at the same time it accounts for more of the specific characteristics of individual observations. At some point, the model is less driven by the general behavior but by "knowledge" on how best to accommodate the noise the data sample. When applied in the real world, the latter rules in the model are detrimental to out-of-sample predictive power. Figure 1 illustrates this effect: focusing on the most relevant factors tends to provide more robust performance. This is also an ecologically rational strategy: the most important factors in an environment have the highest likelihood to stay important (Todd & Gigerenzer, 2000)

Figure I-1: In-and out-of-sample predictive power, as a function of the number of free parameters



Achieving parsimony and simplicity during in-sample model development can be achieved in multiple ways. First we can consciously stop early when adding new predictors which only marginally improve model performance. Second, we can use advanced algorithms for feature selection which penalize too many free parameters (see e.g. Tian et al. (2015), who use the LASSO least absolute shrinkage and selection operator for variable selection). Third, we can use robustness checks during model development to "stress-test" whether an additional complication of a model adds value.

Validity and robustness checks – One straightforward method to check for robustness is to use multiple sub-samples or *cross-validation*. Thereby we "train" the model on one part of the data, and reserve part of the population or sub-periods for validation tests of model performance. This allows detecting performance degradation as illustrated in Figure 1 above. Resampling algorithms such as *bootstrapping* provide an alternative. They simulate the process of obtaining new data samples and allow estimating confidence intervals for model parameters. If a researcher's preferred model parameterization is an overfitted local optimum on the training sample, bootstrapping will show that performance is at the right hand side of the performance distribution. Ideally, we use the final *hold-out* or *out-of-sample* data only once; otherwise, we become guilty of *data-snooping* from testing too many variants of a model. Effectively, this turns an out-of-sample test into in-sample optimization, and we run the risk of building an overfitted model that fails in reality. In practice, it is very difficult to control whether a researcher did indeed use the out-of-sample test only once. Hence, *data-snooping biases* that lead to false positive empirical findings are a widely discussed topic in empirical research on financial markets (see e.g. Marshall et al., 2008; Park & Irwin, 2007). Again, advanced algorithms such as the *Reality Check* procedure by White (2000) and the *Superior Predictive Ability* test by Hansen (2005) provide ways to adjust for biases. They explicitly account for the overstatement of model performance resulting from data-snooping, by adjusting for the fact that many model variants have been tested to come up with one final solution.

The next section of this chapter will discuss the three essays in the context of these criteria and summarize their contributions to their respective fields of research.

2. Essay Contributions to their Fields and Robust Predictive Modeling

Table I summarizes the methods used by the three studies, following the criteria to achieve robust predictive power outlined in the previous section.

Table I-1: Essay Modeling Approaches to achieve Predictive Power and Robustness

Goals:	**Achieving Predictive Power**		**Achieving Robustness**	
Key requirements:	*Correctly-specified structure / prediction algorithm*	*Adapted to task environment*	*Parsimony and simplicity*	*Validity and robustness checks*
E.1 Rating Prediction across Industries	• Linear regression with nested logit sub-models for financial ratios • Heterogeneity score based adjustment for industry effects	• Fundamental: factor selection guided by economic categories and intuition from financial statement analysis • Statistical: outlier protection via logistic factor transformation	• One regression model covering 9 industry sectors with industry-adjusted prediction factors • One procedure for all predictors: Box-Cox, heterogeneity score, logit transformation	• Out-of-sample validation sample • Bootstrap validation of model performances
E.2 Attention-based Trading Strategy	• Adaptive trading heuristic based on regression model	• Fundamental: market capitalization as proxy for "noise trader" impact/profit potential	• Trading heuristic with very few parameters • One regression term driving all entry and exit orders	• Sensitivity test • Out-of-sample validation • Actual-dollar (!) trading
E.3 HFT causes Intraday Extreme Events (IEE)	• Use of exogenous shocks to isolate causal effect of HFT on IEE • Time/firm cluster heteroscedasticity adjusted std. errors	• Fundamental/ statistical: day-time normalized volatility benchmark	• Standard controls encompassing key drivers of volatility	• Cross-validation with sub-periods

2.1. Essay 1 – Improving Performance of Corporate Rating Prediction Models by Reducing Financial Ratio Heterogeneity

Chapter II – co-authored with Jan Hendrik Schmidt and Max Neukirchen – introduces an approach to measure and reduce group-level *financial ratio heterogeneity* in rating models for large multinational corporations. In our case, this heterogeneity stems from industry groups with significantly different balance and income sheet structure: every industry has different typical financial ratio values for a given credit risk category. Due to the limited number of large corporations available for model building, this heterogeneity poses a challenge. One the one hand, the sample is too small to construct separate rating models for individual industry groups; on the other hand, a "one-size-fits-all" model's performance suffers from industry noise in the data.

The heterogeneity reduction approach takes its power from a granular measurement of the industry-driven differences of relationships between financial ratios and credit ratings. We minimize the sample-size weighted average of these measurements with an iterative procedure, by adding or subtracting an adjustment value for each industry group. We show that reducing this definition of financial ratio heterogeneity results in a rating prediction model with better performance than both unadjusted models and models adjusted by including industry dummies or other simpler procedures.

The paper contributes to the literature in several ways: first, the heterogeneity approach fills a gap in cases where a limited dataset does not permit the construction of separate models for individual industries or regions. Second, while previous literature focused mostly on choosing the optimal *classification algorithm*, our paper shows that *factor definition* and *factor transformation* yield further performance increases. Furthermore, this framework of performance levers is a useful tool to help constructing optimal bankruptcy prediction models in general.

Methodologically, the study showcases all criteria to achieve robust predictive power. The chosen *classification algorithm* is well suited for the task. We use linear regression with nested sub-models (e.g. logit and gaussian transformations) for each financial ratio. The linear link function predicts probabilities of defaults that have been linearized with a log-transformation. The second layer allows us to account for different kinds of non-linear relationships between a company's financials and credit risk and at the same time takes care of outliers in financial ratio distributions. This outlier protection *adapts* the model *to the task environment*. For predicting credit risk, it is desirable to make a balanced assessment, analyzing different aspects of a company's financials. Capping outliers ensures that no single variable takes too much weight in the regression results. Our approach to factor selection also contributes to this goal: we perform an economic categorization of financial ratios and make sure that the most important categories are covered. In addition, we reduce multi-collinearity by barring too many similar variables from entering the model. Together, all these techniques and methodology choices help achieving *robustness*. We are able to solve a challenging multi-industry prediction problem with a comparatively *parsimonious* regression model. The factor transformations – Box-Cox, heterogeneity reduction, and construction of non-linear sub-models – are uniformly applied. We employ multiple *checks of validity and robustness*. For all model candidates, we train the algorithm on a training sample, which comprises 80% randomly chosen observations, and then compare model performances *out-of-sample* on the remaining 20%. Furthermore, with a bootstrapping test, we validate the performance improvement of the heterogeneity-score enhanced model versus alternatives from the literature.

2.2. Essay 2 – Exploiting Attention-Driven Mispricing: Evidence from Actual-Dollar Trading

This study – presented in full length in Chapter III – performs a real-life test of the *efficient market hypothesis* by Fama (1970), which suggests that it should be pointless to design a statistical trading strategy based on past data and expect to earn *excess returns* beyond a fair compensation for the risks of deviating from the market portfolio. We do exactly that. Not only does the study show that we can earn statistically significant trading profits, it does so *out-of-sample* with *real money*. Making real trades circumvents all needs to make assumptions about market frictions such as trading costs and short-selling constraints, which make up a significant part of the academic discussion around whether theoretical trading strategies can actually make economically significant profits (see e.g. Malkiel, 2003).

The trading strategy sells short U.S. stocks at market open on the day following a buy recommendation by Jim Cramer in the evening TV show ``Mad Money'' on CNBC. It exploits the published phenomenon of attention-driven buying by retail investors, whereby stock recommendations cause mispricings by drawing investor attention to stocks, although the recommendations themselves have zero information value. Hence, the initial price spike from a recommendation is typically reversed in full over several days. The trading strategy covers most positions profitably, with an average 0.53% return per trade, in the first half-hour of the day. Adjusting for common risk factors in the literature does not explain the excess returns.

Further investigations into the source of the strategy's profits show that intraday, recommended stocks exhibit a ``reversal-of-the-reversal'' pattern, i.e. after an initial correction of the mispricing, it re-widens temporarily. This intraday phenomenon is undocumented in the literature on attention and recommendation effects, and supports theories on limits to arbitrage and destabilizing speculation. The strategy captures these intraday dynamics better than an average market participant, suggesting excess returns from market timing. This constitutes out-of-sample evidence of market inefficiency, though by itself not economically large. Nevertheless, if the observed intraday patterns in stock returns were universal, they would represent a significant deviation from market efficiency.

The strategy is also an example of *robust predictive modeling* used for real-life decision-making. The classification algorithm is a simple trading heuristic – with an entry rule, and a profit-target and stop-loss rule for exit – that uses one simple regression of mispricing size vs. stock size, measured by market capitalization. This simple feature, however, adapts the strategy very well to the task of predicting the price impact of retail traders who react to stock recommendations: the smaller a stock, the higher the market share that these "noise traders" can take, and the higher the mispricing and profit potential for the trading strategy. The trading model is simple, with very few free parameters. After estimating the size of the mispricing as a function of company size, this prediction drives entry, profit-target and stop-loss strategies. For *validation and robustness checks*, we use two *out-of-sample* periods – one before and one after the training period. Having confirmed the model's predictive power in out-of-sample simulation, we apply it for real-life trading decisions. What form of validation could be more powerful for a model than using it successfully in reality?

2.3. Essay 3 – High Frequency Trading Intensifies Intraday Extreme Events in Stock Returns

Chapter IV presents an investigation of the impact of *high frequency trading (HFT)* on intraday extreme events in U.S. stock returns. The influence of HFT on market characteristics, and whether HFT increases or decreases *market quality*, are disputed questions in the literature. Published empirical studies have typically found that, on average, HFT improves standard measures of market quality, such as average spread, volatility, and short-term autocorrelation of returns. On the other hand, the "Flash Crash" of May 6, 2010 is an event for which empirical investigations show that HFT has intensified selling pressure. Furthermore, numerous theoretical models as well as observations by market practitioners suggest mechanisms by which HFT could systematically cause or contribute to large short-term price moves, or *intraday extreme events (IEE),* in stock returns. We test for a causal relationship between HFT and IEE with a multi-year sample of 1-minute price and HFT activity data on all liquid U.S. stocks. The IEE measure effectively measures tail risk in the form of "X-sigma events", comparing intraday price moves with a measure of a stock's daytime-adjusted typical volatility. We isolate the causal effect of HFT on IEE with an instrumental variable regression, using exogenous shocks to HFT activity as instruments. Regulation NMS (summer 2007) acts as a positive shock to HFT activity, whereas the SEC Naked Access Ban (winter 2011/2012) constitutes a negative shock. Across both sub-periods, we find that HFT activity exacerbates intraday extreme events with statistical and economic significance. These results add empirical evidence to the debate among researchers, market participants and regulators about the benefits and drawbacks of HFT. Currently, investors "pay" for the benefits of HFT market participation by having to bear larger short-term tail risks in stock returns. For non-HFT investors, this increases adverse-selection and hedging costs.

This paper closely follows the framework of *predictive and robust modeling*. Using an instrumental variable approach in a panel regression framework disentangles the endogeneity of HFT. Are high frequency traders active because the market is volatile, or do they cause the volatility? Using instrumental variables is also a risky approach. Numerous assumptions have to hold for results to be viable. Since we cannot be 100% sure that our assumptions are correct, we use two sub-periods with two different instruments for cross-validation. Another risk in the panel regression is heteroscedasticity across firms and time, which we address by using double-clustered standard errors in all regressions. The measure of intraday extreme events incorporates a lot of our previous knowledge about intraday volatility: it is highly different across stocks and daytime. Calculating a stock- and daytime-specific benchmark of what constitutes "extreme", we obtain an IEE measure that is highly consistent across stock size classes and time. These are desirable properties for our task of investigating the marginal impact of HFT activity. Next to HFT and IEE, the regressions stay very simple, using only standard control variables from the HFT literature. Apart from control variables, we test for robustness with specification changes and through the above-mentioned cross-validation across multiple periods. This comes as close as we can get to a true out-of-sample test.

3. Overarching Learnings on Financial Market Behavior

Beyond individual contributions to respective literatures and the application of robust predictive modeling principles, the three studies yield several overarching insights on the laws governing financial markets: the importance of *firm size*, the existence of *pockets of market inefficiency* and the fact that *agent interaction effects can trump fundamentals* in determining asset prices, at least in the short-term.

The importance of firm size is universal across all three prediction problems and their respective literatures. In essay 1, firm size, measured by the book value of equity or assets, enters *every* credit rating model with the highest weight. This makes sense as size proxies for diversification and economic resilience. A large firm which is active in multiple products, markets, and geographies with a large and diverse asset base can withstand economic shocks much more than a small firm with little reserves and concentrated income streams. Numerous other sources (e.g. Režňáková & Karas, 2014; Tian et al., 2015) in the rating literature confirm this finding. In essay 2, firm size – this time measured by market capitalization – influences *every* trading decision in the strategy from entry to exit. Again, this makes sense given what we know about asset pricing: size is one of three factors in the seminal paper by Fama & French (1992). In our case size proxies inversely for the relative market share which retail noise traders can achieve in a stock, and thereby drives our profit potential. In essay 3, size enters the arena again in the form of market capitalization. Along with size, measures of volatility, liquidity and trading volume vary systematically. Consequently, so differ the effects of HFT that we seek to measure. We adopt the approach from prior studies, cutting the universe of stocks into size terciles (small, medium, and large), and confirm substantial differences in HFT activity by firm size.

Pockets of market inefficiency exist in financial markets are partly linked to the third notion, that ***agent interactions can trump fundamentals***. In essay 2, we show that retail noise traders create a profit opportunity that is systematically exploitable, with economically significant profits. Theories of agent interaction – in this case, between *noise traders* following positive-feedback strategies, and *arbitrageurs* who attempt to profit from an apparent mispricing – provide a potential explanation why the pattern remains exploitable for a longer period. In addition, we find a "reversal-of-the-reversal" pattern, i.e. a temporary re-widening of the mispricing. This supports theoretical models of destabilizing speculation by De Long et al. (1990) and underlines the need for arbitrageurs to synchronize their market-impacting trades to successfully correct mispricings. (Abreu & Brunnermeier, 2002). When arbitrageurs fail, positive feedback traders can drive prices further away from fundamentals, forcing arbitrageurs to liquidate their positions at a loss, whereas in the long-term, the market would reward the opposite transaction. In such a trading environment, it can be a rational strategy to trade *against* fundamentals, leading to a re-widening the mispricing. This is exactly what the trading strategy in essay 2 does by existing its short-sales early – it buys back stocks although the mispricing has only been partially corrected. Although contrary to fundamentals, the action is rational as it preempts the same move from other arbitrageurs. What we learn from the trading strategy's profitability is that these agent interactions create at least a pocket of exploitable market inefficiency. Patterns created by interaction of short-term traders are stable enough to construct a model for market timing and apply it successfully in the market. Further research would be required to show that this is universal.

Essay 3 also supports the view that ***agent interaction effects can trump fundamentals*** in the short-term. It shows that there are short-term market phases in which high frequency traders contribute to outsized short-term price moves. While we cannot differentiate whether these intraday extreme events are created by *one* HFT, the interaction of several HFT or the interaction of HFT and non-HFT participants, one aspect stays constant: highly sophisticated players in the market employ strategies driving prices away from fundamentals. In combination, essays 2 and 3 suggest that several mechanisms exist that can drive asset prices away from fair value – sometimes the source are retail traders, sometimes institutions operating at the margin of technical and analytical sophistication. As shown in essay 2, some of these phenomena give rise to trading opportunities which are economically and statistically significant, and hence, constitute instances of market inefficiency.

II. Essay 1 – Improving Performance of Corporate Rating Prediction Models by Reducing Financial Ratio Heterogeneity

Authors

Martin Niemann (primary author, 50% authorship)
University of Augsburg, Faculty of Business and Economics, Institute of Statistics
Universitätsstrasse 16. 86159 Augsburg, Germany
martin.niemann@gmail.com

Dr. Jan Hendrik Schmidt (25% authorship)
McKinsey & Company
janhschmidt@hotmail.com

Dr. Max Neukirchen (25% authorship)
McKinsey & Company
max.neukirchen@gmail.com

Status

Published in
Journal of Banking & Finance (2008), Vol. 32, 434-446
DOI: 10.1016/j.jbankfin.2007.05.015

Abstract

We introduce a new approach to improve the performance of rating prediction models for multinational corporations. In this segment, the low number of defaults poses a challenge, as it prevents rating models to be constructed for individual industry sectors or regions. We show that reducing group-level heterogeneity in financial ratios results in a rating prediction model with better performance than both unadjusted models and models adjusted by including industry dummies or other simpler procedures. Our approach fills a gap in cases where a limited dataset does not permit the construction of separate models for individual industries or regions.

III. Essay 2 – Exploiting Attention-Driven Mispricing: Evidence from Actual-Dollar Trading

Author

Martin Niemann
University of Augsburg, Faculty of Business and Economics, Institute of Statistics
Universitätsstrasse 16. 86159 Augsburg, Germany
martin.niemann@gmail.com

Status

Presented on Jan 29, 2015
at the Institute of Statistics, University of Augsburg, to Prof. Dr. Yarema Okhrin and Prof. Dr. Dr. h.c. Bamberg.

Abstract

This study trades actual dollars in U.S. stocks during the day following a buy recommendation by Jim Cramer in the evening TV show “Mad Money” on CNBC. Prior papers show that Cramer's stock picks have no information value, yet induce attention-driven buying by retail investors, causing mispricings which disappear over several weeks. The trading strategy sells short stocks at market opening after the recommendation and covers most positions profitably in the first half-hour of the day. Profits are statistically significant and average 0.53% per trade. Common risk factors do not explain the returns. Intraday, we find that the mispricing is not corrected by a monotonous price decline as suggested by previous literature on recommendation effects. Instead, we find that the initial down-move is reversed by a significant transitory *increase*. This re-widening of mispricing supports theories on limits to arbitrage and destabilizing speculation. The strategy captures this intraday pattern better than an average market participant, suggesting excess returns from market timing. However, due to its small size, this is a small deviation from market efficiency. In sum, the business of exploiting mispricing seems risky and competitive.

1. Introduction

This study started in an unusual way for a financial economics paper. It started as a practical project with the primary goal to make money – by developing a statistical arbitrage strategy and trading it with real capital. However, beyond greed, there was an underlying research motivation relating to financial economics: it was to disprove friends, co-workers and the majority of market efficiency literature arguing that the market is so efficient that it should be pointless to try to predict future stock returns based on past information which should already be reflected in market prices (Fama, 1970). It should be pointless to design a statistical trading strategy based on past data and expect to earn excess returns beyond a fair compensation for risks such as investing in small firms (see e.g. Chan & Chen, 1988). Hence the practical project did have a relevant underlying research question:

> Can an *ex-ante*-specified trading rule earn *excess* returns in an *out-of-sample test* using *actual dollars* in the stock market?

We start by summarizing the practical project of developing and trading the statistical arbitrage strategy. Then we investigate the results and observations and link them to existing literature. The trading strategy builds on a working paper by Engelberg et al. (2006) (hereafter ESW), who study the reactions of stocks after recommendations by Jim Cramer in the popular evening TV show "Mad Money" on CNBC (hereafter *Cramer events*). They find a significant "spike-reversal pattern" with abnormal returns on the day following the recommendation – 5.19% for the smallest quartile of stocks – which disappear within 12 trading days. ESW also find that statistical arbitrageurs respond, with high short-sales volume in the opening minutes following the recommendations. However, without data on stock loan fees and transaction costs, they do not conclude whether significant predictability also allows earning excess returns.

ESW's findings motivated investigating Cramer events with the aim of earning real profits. We use data on prior Cramer events and daily Open, High, Low, Close prices and Volume of recommended stocks from July 2005 to October 2006. We predict the overnight price impact to sell short attractive events at the Open price on the day following the recommendation (hereafter *trading day*). The strategy closes (buys to cover) its short positions on the same day, either via a profit-taking limit order (below the entry price), a trailing stop-buy order (initially above the entry price, then moving down with price) or at market close. With a simulated Sharpe ratio of 3.05 in the *in-sample* period and successful *robustness tests* across subperiods, our group of investors decided to put real money at stake. During the live trading period from December 2006 to July 2007, the strategy has entered 68 short positions with an average size of approximately US-$174,000, earning a cumulative unleveraged net return of 36.3% with a Sharpe ratio of 3.30 (see Table 2). Clearly, actual-dollar trades confirm ex-ante simulation. Against the null-hypothesis of not trading, the average profit per trade of 0.53% is significant at the .001 level.

These results show significant, economically exploitable return predictability, generated *out-of-sample* with an *ex-ante*-specified trading rule. However, do these profits also constitute *excess returns*? To investigate this, we regress the strategy's weekly returns against contemporaneous returns for main risk factors in the market efficiency literature. We use the results in Bolster & Trahan (2009) for validation, who find that Jim Cramer's picks have exposure to Size, Value and Momentum, and high market betas. Since we *short* the recommendations we should find *negative* coefficients on these risk factors. We find exactly that. However, we do find a significantly positive exposure to Reversal. This suggests that part of the strategy's profits are a compensation for providing liquidity (Pástor & Stambaugh, 2003). Still, the share of profits explained is very small. Unexplained "Alpha" is the most significant coefficient by far.

There are two possible interpretations of these unexplained profits: (1) the strategy indeed managed to earn excess returns, (2) a regression on daily or weekly returns is simply a too blunt tool to explain profits made intraday, which otherwise *could* be explained by risks. Unfortunately, we are not aware of a model of intraday expected returns in the literature. Nevertheless, we do know that the strategy is somewhat risky, since part of the strategy's trades end with losses. To disentangle whether the strategy earned its profit from taking risks or from above-average market timing ("Alpha"), we analyze how the strategy navigates the intraday price paths of stocks. The strategy closes more than half of its trades during the first 20 minutes of the trading day, and more than two thirds in the first half-hour. Moreover, these short-term trades are no less profitable than their longer-term counterparts. There is a bias towards triggering the profit target or the trailing stop-loss order very early in the day. This implies that stocks failing to fall enough to reach the profit target tend to *reverse upwards* to trigger the trailing stop-loss, rather than settling at some "fair value" price until market close. We test for this *"reversal-of-the-reversal"* pattern with intraday regressions on Cramer events and confirm its existence. As expected, for the first half-hour of the trading day, we find that the size of the initial downtrend is associated with the initial mispricing. However, in the second half-hour, there is a significant tendency for the initial *down*-move to be reversed in a transitory *up*-move. In other words, there is a significant tendency for a *re-widening of mispricing* intraday.

The strategy navigated this complex pattern better than the average market participant. The expected profits of the assumed *average* arbitrageur on the first post-event day amount to about 33.5% earned with a holding period of six and a half hours for the 68 observed trades. The strategy however, earns 36.3% with an average holding period of only 34 minutes. The strategy extracts additional profits from the intraday price path of stocks and thereby outperforms the average arbitrageur. This suggests that next to a premium for short-term liquidity provision, the strategy also earned excess returns from market timing.

Taken together, our findings relate and contribute to several bodies of existing literature: (1) tests of market efficiency, (2) impact of stock recommendations and attention, and (3) limits to arbitrage.

The first literature of interest comprises tests of *market efficiency*, *return predictability*, and the profitability of *trading rules*. Over time, the literature has developed at least four challenges, or tests, which a potential new "market anomaly" has to pass: (i) controlling for *data-snooping bias*, (ii) accounting for *market frictions* and (iii) *micro-structure effects*, and (iv) adjusting for known *risk factors*. An actual-dollar out-of-sample test is rare in the literature[1] and conveniently addresses the first three challenges. *Ex-post* studies suffer from the risk of data-snooping, i.e. that researchers incorporate their knowledge of market history in their choice of prediction model (Timmermann & Granger, 2004; Avramov & Chordia, 2006; Cooper et al., 2005). A real-money out-of-sample test rules out this risk. The second and third challenges, micro-structure effects such as non-synchronous trading and bid-ask-

[1] We are not aware of any comparable study setup. Seasholes & Wu (2007) come close by observing other arbitrageurs profiting by trading against retail traders.

bounce (Amini et al., 2013) and market frictions such as trading costs and short-selling constraints (Malkiel, 2003), are solved by actual-dollar trading as well: there is no risk of overly optimistic assumptions. Finally, our test of common risk factors finds that they explain only a small part of profits. With most risk factors in the literature ranging between 2 - 11% risk premium per annum (Asness et al., 2014), high intraday returns warrant further research as to their source.

The second literature of interest studies *recommendation effects* and the impact of *investor attention* on asset prices. The reversal-of-the-reversal pattern is a new finding to the literature on recommendation effects[2] as well as the body of papers on Cramer events in particular, such as ESW and others[3]. These studies consistently find abnormal returns which subside to zero after a few days, weeks or months. By and large, stock recommendations have no information value. Their only impact is a temporary mispricing. In the published version of their paper, Engelberg et al. (2012) causally link retail investors' attention to the size of the mispricing after Cramer events, confirming empirically that retail traders are attention-driven "noise traders" (Barber & Odean, 2008). Thus, what noise traders pay as a price premium for following the recommendations provides the incentive for smart arbitrageurs to provide liquidity. Our results suggest two extensions to this view. First, our analysis of the strategy's market timing implies *competitive pressure* among arbitrageurs. In our example, in addition to naïve investors, even the average arbitrageur paid part of our profits. Realizing our profits by buying back our short position early in the day

[2] see e.g. Kerl & Walter (2007), Busse & Green (2002), Liang (1999), Metcalf & Malkiel (1994), Wright (1994), and Barber & Loeffler (1993)

[3] see e.g. Bolster et al. (2012), Hobbs et al. (2012), Chen et al. (2011), Keasler & McNeil (2010), Lim & Rosario (2010), Bolster & Trahan (2009), Karniouchina et al. (2009), Neumann & Kenny (2007) and Engelberg et al. (2006). Going beyond the recommendation literature, there is a somewhat comparable finding by Kudryavtsev (2013), who find that intraday reversals on one day are reversed *again* on the following day. However, the average reversal found is only 7-8 basis points per day.

implies that later trades by other arbitrageurs occur – on average – at an inferior price and with less efficient use of capital. Second, the intraday re-widening of the mispricing increases the risks for arbitrageurs and indicates that it could be rational to speculate in the original direction of the mispricing, rather than against it.

This notion leads us to the third literature of interest on *limits to arbitrage*, which investigates factors that deter arbitrageurs from correcting mispricings. In contrast to the theoretical arbitrageur who has unlimited capital to trade against noise traders, real-life arbitrageurs do not (Shleifer & Vishny, 1997). If a price shock from noise traders is too large, arbitrageurs might fail to correct mispricings and be forced to liquidate their positions at a loss. This risk might deter arbitrageurs from trading against the mispricing. Abreu & Brunnermeier (2002) suggest that arbitrageurs have to join forces to correct a mispricing, and thereby incur *"synchronization risk"*. As a result, successful arbitrage requires timing the market, including the actions of other arbitrageurs. Our observations are consistent with these predictions. Many actual-dollar trades were exited via stop-loss as additional demand increased the mispricing beyond the risk limit set by our group of investors (i.e. arbitrageurs failed to synchronize). Furthermore, if noise traders follow positive-feedback strategies and/or other arbitrageurs might be forced to liquidate, De Long et al. (1990) suggest it is rational to jump on the bandwagon, driving prices even further away from fundamentals. Our finding of a systematic tendency of the mispricing to re-widen intraday supports this.

Next to its profitability, our trading experiment shows that the "business" of correcting mispricing is highly competitive and risky.

The outline of this paper as follows. Section 2 describes the dataset of Cramer events used in section 3 to develop the statistical arbitrage strategy. Section 4 presents the actual trading results and further investigations as to the source of profits. Section 5 concludes.

2. Data

The development of the statistical arbitrage strategy uses a sample of 983 Cramer events from July 28, 2005 to October 17, 2006 collected by the independent website YourMoneyWatch.com.[4] For each recommendation, we gather historical stock prices (daily *open*, *high*, *low* and *close* prices, and *volume*) for the respective trading days $t-30$ to $t+15$ around the recommendation day t, as well as stocks' market capitalization from Yahoo! Finance. We remove 67 due to missing data, leaving a sample of 916 Cramer events for analysis. We split the sample to obtain out-of-sample data for model validation. We chose the 393 Cramer events from January to July 2006 as in-sample dataset, leaving two out-of-sample datasets for analysis, with "out-of-sample I" containing 106 events from August 1, 2006 to October 17, 2006, and "out-of-sample II" comprising the remaining 417 observations from 2005.[5] Descriptive statistics in Table 1 show that across periods, stock characteristics such as market capitalization and liquidity differ somewhat across subperiods, whereas price

4 While recommendations are also tracked on Jim Cramer's affiliated website TheStreet.com, Engelberg et al. (2006) find that YourMoneyWatch.com applies higher standards in distinguishing what constitutes a "buy" recommendation versus merely positive mentions of a stock during the show. For instance, unconditional buy recommendations such as "this stock is a triple buy" are considered a buy, whereas conditional statements such as "I like the stock, but wait until it drops below $60" are excluded.

5 From the perspective of a financial economist, the methodology used in the practical trading experiment is far from optimal. Developing a strategy for a group of wealthy individuals required some pragmatic short-cuts. The periods used for strategy development and trading are non-standard, i.e. not of uniform length, such as years or half-years, as they were driven by the time when data became available and the opportunity to present the strategy to several wealthy individuals and obtain capital to trade. As a result, actual trading does not use a constant amount of capital, as the investors increased their amount invested after the first profitable trades had increased their confidence in the strategy. Trades are overlapping in time and hence should not be treated as independent observations, since contemporaneous returns for the market or other risk factors could drive results. Nevertheless, the benefit of a realistic, unbiased out-of-sample test are worth a few methodological sacrifices.

and volatility are largely similar. Since the trading strategy uses logarithms of market capitalization in its formulas, these differences have little impact.

Table III-1: Descriptive statistics for in-sample, "out-of-sample I" and "out-of-sample II" data

Notes: This table shows descriptive statistics for Cramer events (clear "buy" recommendations by Jim Cramer as documented by YourMoneyWatch.com) for stocks on the recommendation date t0, with the actual recommendation occurring later during the evening of the same day. All data have been provided by Yahoo! Finance. Volatility is calculated as the standard deviation of daily log. returns from t-19 to t0. Liquidity is defined as price times volume at the close of t0.

		out-of-sample I	in-sample	out-of-sample II
N (Cramer events)		106	393	417
First/last recommendation date		Aug. 1, 2006 - Oct. 17, 2006	Jan. 3, 2006 - Jul. 31, 2006	Jul. 28, 2005 - Dec. 23, 2005
Closing price at t0* (US-$)	Average	38.45	32.82	38.19
	Median	34.06	28.60	32.65
	Standard deviation	22.19	21.53	29.46
	Min	4.03	2.39	1.67
	Max	117.65	131.04	356.51
Market capitaliza-tion at t0* (US-$ millions)	Average	23,672	12,929	21,478
	Median	6,435	2,020	6,670
	Standard deviation	40,868	30,681	43,086
	Min	252	93	57
	Max	284,890	248,490	431,740
Volatility t-19 to t-0 (percent)	Average	2.0	2.3	2.0
	Median	1.7	2.1	1.8
	Standard deviation	0.9	1.2	1.0
	Min	0.6	0.5	0.6
	Max	5.4	8.3	8.7
Liquidity t0* (US-$ millions per day)	Average	238.7	90.1	102.1
	Median	47.2	23.1	36.3
	Standard deviation	527.0	263.2	198.8
	Min	0.4	0.1	0.1
	Max	2,706.8	3,920.4	2,287.3

* closing price on recommendation day (recommendation after market close)

3. Development of the Trading Strategy

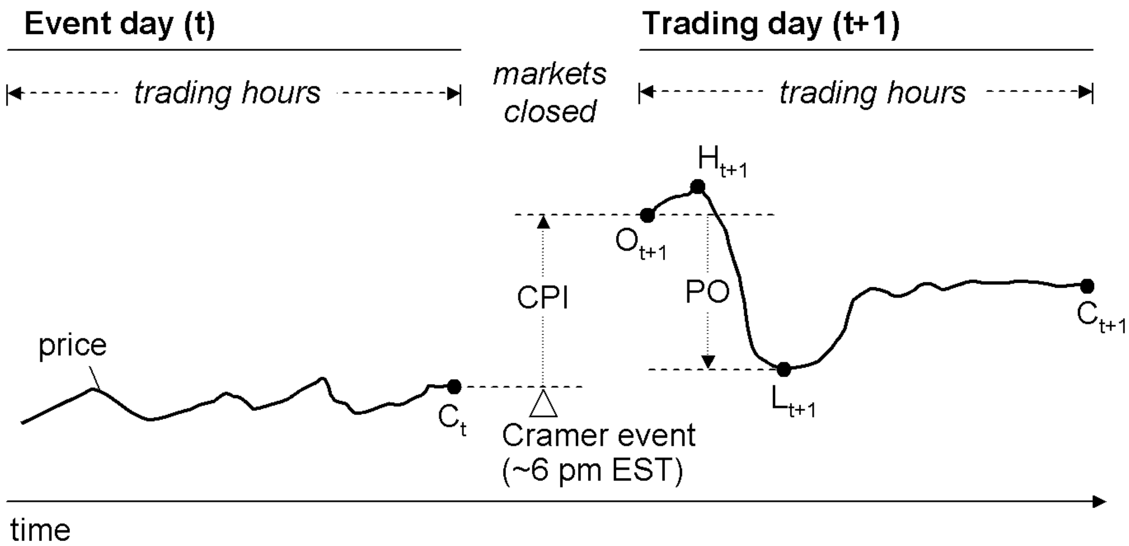
This section describes the strategy development performed before the start of actual trading on December 4, 2006. We start by analyzing the profit opportunity from Cramer events to derive the requirements for trading strategy design. The second subsection describes the strategy and its parameters, followed by the in-sample and out-of-sample simulation results in the last subsection.

3.1. Profit Opportunity in Cramer Events

Figure 1 provides a schematic overview of the sequence of events and relevant variables used for modeling the trading strategy over the course of the event day t and the following day $t+1$. The indicated price path shows the stylized reaction observable in stock prices after a Cramer event. The "Mad Money" show airs every weekday at 6 p.m. Eastern Standard Time, two hours after market close, during which Jim Cramer features one to five stocks per day as explicit "buy" recommendations. The results in ESW as well as other studies on recommendation effects (see above) show that these stocks reopen at a significantly higher price in $t+1$, as a result of orders entered by Cramer's audience while the market was closed.

Figure III-1: Sequence of events and relevant data points for modeling

Notes: This figure shows a schematic overview of the sequence of events and relevant variables used for modeling the trading strategy. The indicated price path provides a typical example of stocks' reaction to Cramer events. CPI denotes the overnight price impact – defined as the opening (O_{t+1}) to prior close (C_t) return – of Cramer's "Buy" recommendations, whereas PO represents the maximum theoretical profit opportunity to a trader who shorts stocks at the open, and manages to cover the position at the day's low price (L_{t+1}).



Therefore, the strategy hypothesis is to sell short at the opening price in $t+1$ in the expectation of buying the stock back at a lower price once abnormal returns subside. Given the daily flow of new events, the strategy should exit positions no later than at the close of $t+1$ to free up capital for incoming Cramer events on the following day. The historical daily opening (O), high (H), low (L) and closing (C) prices provide the vertices of a stock's intraday price path; however, since the sequence of H and L is unknown, testable strategies have to be based on conservative assumptions. As a consequence, the only sure entry point for a short-sale is at O_{t+1}.

Two observable variables are of particular interest for understanding the structure of the profit opportunity, as shown in Figure 1:

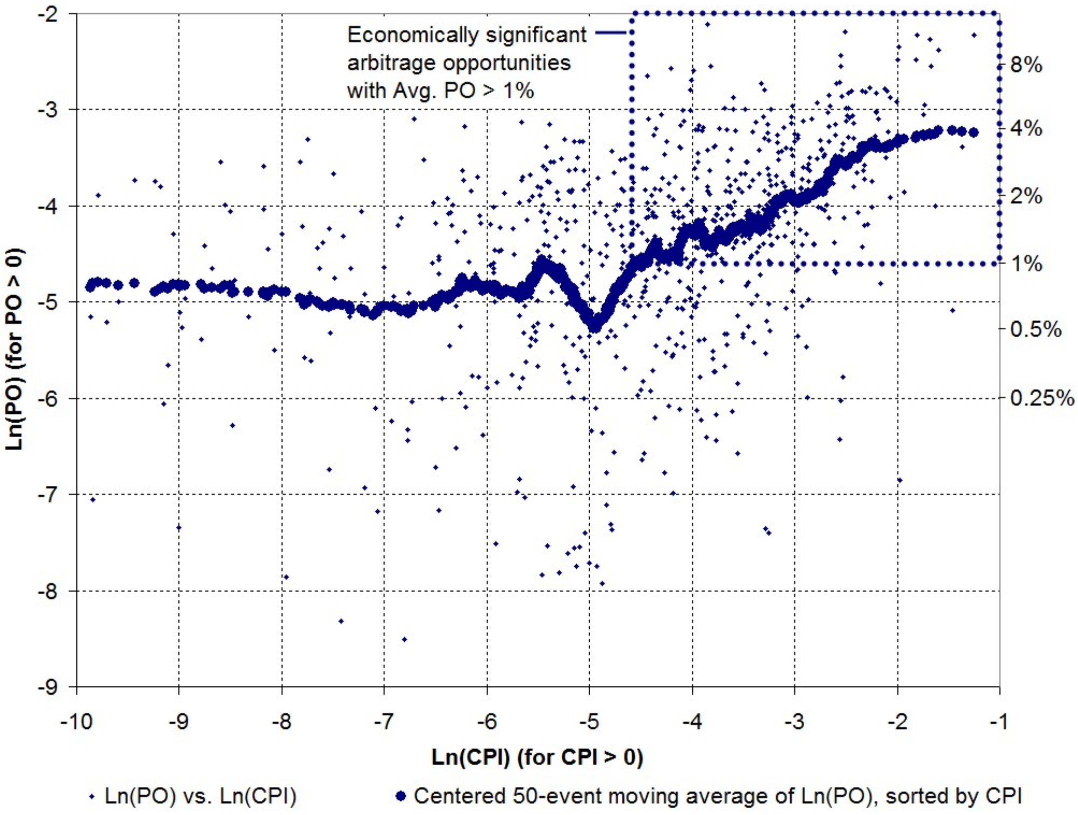
(1) The initial price impact caused by Cramer's recommendation (further referred to as *Cramer price impact, or CPI*) is defined as the close-to-open return $O_{t+1}/C_t - 1$.

(2) The *profit opportunity (PO)* is defined as $O_{t+1}/L_{t+1} - 1$, and indicates the maximum profit attainable for a short-sale at the opening of the trading day, when achieving the optimal exit price.

Results of a study by Zawadowski et al. (2006), who find a significant correlation between the size of intraday price spikes and the extent of subsequent reversals, lead us to believe that PO should be a function of CPI. For the purpose of short-selling, this only makes sense if a Cramer event actually has a price impact greater than zero. Figure 2 shows the logarithmized scatter plot for PO as a function of observed positive CPI values. Indeed, CPI and PO are correlated. This allows isolating the events with attractive profit potential. For events with observed CPI below 1.0% (corresponding natural logarithm of approximately -4.6), PO becomes economically unattractive to trade.

Figure III-2: Logarithms of profit opportunity (PO) vs. Cramer price impact (CPI), and economic "priority area" for trading

Notes: The scatterplot is based on 741 Cramer events with $CPI > 0$ and $PO > 0$ from July 28, 2005, to October 17, 2006. The dotted box highlights events of higher economic significance (approx. 450 events), for which expected $PO > 1.0\%$, as given by the centered 50-event moving average of Ln(PO) values. The %-values on the secondary Y-axis indicate corresponding delogarithmized values for Ln(PO). This graph is shown "as-is" from the strategy development process.



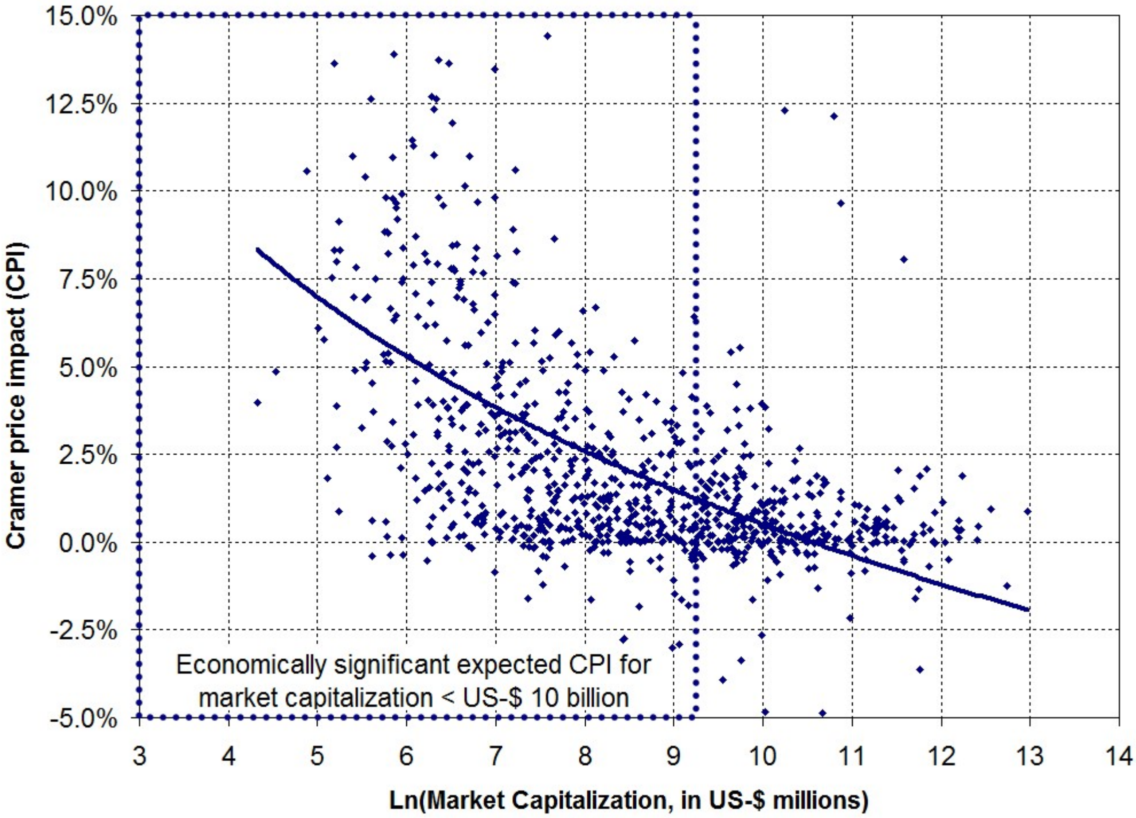
Engelberg et al. (2006) show that the impact of Cramer events varies substantially with firm size. Hence, we regress CPI on stocks' market capitalization (M) to obtain an indication of *expected* CPI, as shown in Figure 3. As expected, small-capitalization stocks react more strongly to a Cramer recommendation than large capitalization stocks. Again, we prioritize economically attractive events by excluding stocks with a market capitalization above US-$10 billion, corresponding roughly to an expected CPI of 1.0%. A simple logic to use these observations for an entry filter would be to short only those stocks with market capitalization $\leq$ US-$10 billion for which observed CPI reaches at least its expected value CPI_e, as given by the regression function shown in Figure 3.

$$CPI_{ei,t+1} = 0.2277 - 0.0966 * Ln(Ln(M_{i,t})), \qquad (1)$$

whereby $M_{i,t}$ is a stock's market capitalization (in million US-$) for firm i on day t, and CPI_e indicates the expected Cramer price impact in percent.[6] This entry filter selects 209 of all 916 Cramer events for trading.

Figure III-3: Cramer price impact (CPI) vs. logarithm of market capitalization, in US-$ millions

Notes: Scatterplot is based on 916 Cramer events with available market capitalization data, from July 28, 2005, to October 17, 2006. CPI is defined as $O_{t+1}/C_t - 1$. The log-linear regression function is given in Eq. (1). The dotted box highlights selected events with higher economic attractiveness, for which market capitalization ≤ US-$10 billion. This graph is shown "as-is" from the strategy development process.



The next subsection discusses the development of the full trading rule and the simulation procedure to optimize its parameters.

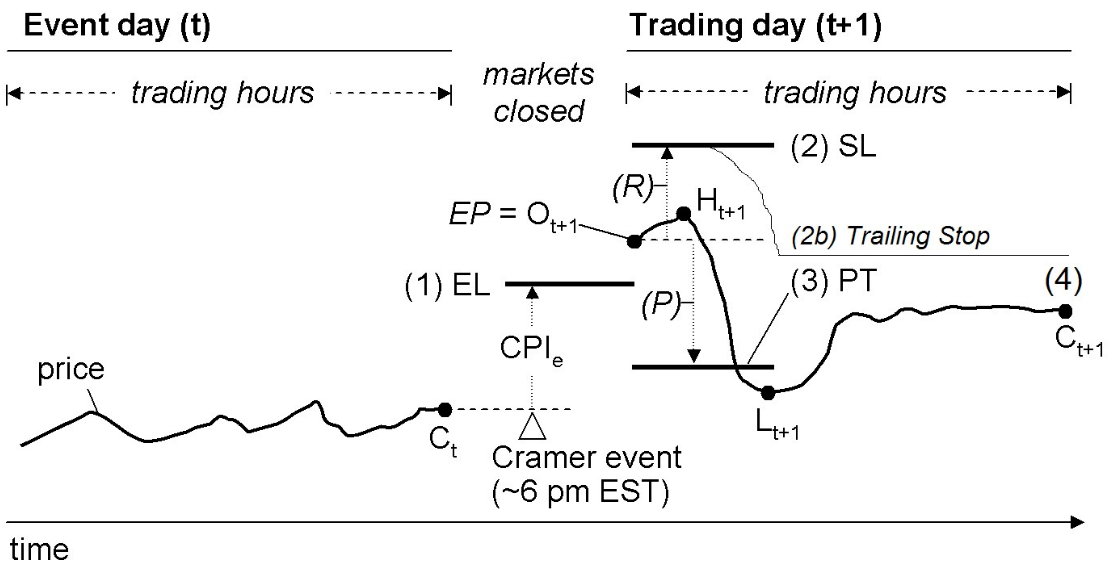
6 The double logarithm of M was chosen since CPI vs. $Ln(M)$ is still notably non-linear (see figure 3).

3.2. Trading Strategy Design

A key requirement for the trading strategy was that it remained easy to operate when trading in a real brokerage account. Therefore we define a simple framework consisting of a maximum of four orders placed for each Cramer event in $t+1$ (further referred to as "trading day"). We enter positions with a sell limit order (1) for execution at market opening. If filled, we place a stop-buy order (2) as protection against large losses[7], and a buy-limit order (3) for profit-taking. If neither of the two exit limits is triggered, we cover the position with a market order (4) at market close. Figure 4 provides a schematic overview of the different orders and price levels.

Figure III-4: Overview of strategy entry and exit orders, and simulation variables

Notes: This figure illustrates the main variables of the trading strategy. Based on Cramer event data, an estimate CPI_e *of the Cramer price impact is calculated. If the following opening price* O_{t+1} *is at or above the entry limit* EL*, then a short position is assumed with* $EP = O_{t+1}$*. Based on the risk parameter* R*, the stop-loss limit* SL *is placed some distance above* EP *(in live trading, this was implemented using a trailing stop order, which follows prices downwards to protect profits). The profit-taking limit* PT *is set below* EP *according to profit parameter* P*. If neither* SL *or* PT *are triggered during the trading day, the position is closed at* C_{t+1}*.*



The limit prices for orders (1) to (3) are given by the following equations:

[7] In live trading, we used the trailing stop functionality offered by the broker, indicated as (2b) in Figure 4.

(1) The entry limit price EL for the short-sale at the opening is

$$EL_{i,t+1} = C_{i,t} * \left(1 + CPI_{ei,t+1}\right), \tag{2}$$

which is assumed to be executed at a realized entry price (EP) of $EP = O_{t+1}$ if $O_{t+1} \geq EL_{t+1}$ during the backtest simulation.

(2) The stop-loss limit level SL is placed at

$$SL_{i,t+1} = EP_{i,t+1} * \left(1 + R * \frac{CPI_{ei,t+1}}{0.03}\right), \tag{3}$$

and it can be interpreted as a volatility-driven risk limit, using CPI_e as volatility estimate, whereby R is a parameter for the desired level of risk, e.g. 1.5% on invested capital.[8] In the simulation, if $H_{t+1} \geq SL$ we assume to exit the position at SL plus transaction costs.

(3) The profit-taking limit level PT is given by

$$PT_{i,t+1} = EP_{i,t+1} * \left(1 - P - 0.5 * \left(CPI_{ei,t+1} - 0.03\right)\right), \tag{4}$$

which is conceptually comparable to the stop-loss level, whereby P sets the distance to the entry price. During simulation, if $L_{t+1} < PT$, we assume to cover the position at PT plus transaction costs, but only if not $H_{t+1} \geq SL$. This may lead to an assumed exit via stop-loss, although in reality the profit-taking limit might have been triggered first. However, this conservative assumption is necessary since the sequence of H and L is unknown for daily price data.

The resulting strategy has only two parameters for further optimization. R sets the risk distance for the stop-loss limit, and P determines the distance of the

[8] The 0.03 constant was used in discussions with investors, to express R as a multiple of 3% typical volatility, and is reported here without changes.

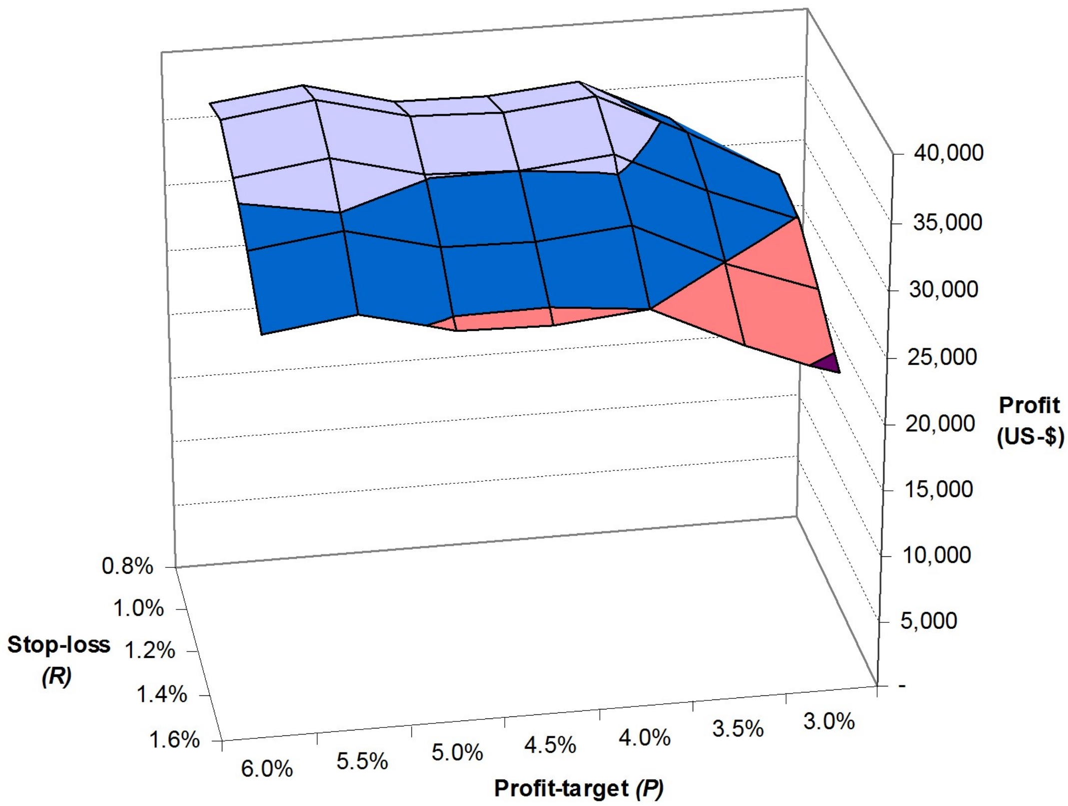
profit-taking limit order. The next subsection presents the in-sample optimization and simulated results for the final trading rule.

3.3. Simulation Results

For the model simulation, we assume to sell short US-$50,000 in stock for every Cramer event triggering the entry limit, i.e., profits are not reinvested. We assume US-$20 broker commission per transaction and add (substract) 0.1% slippage to (from) the price of every buy (sell) transaction to account for potential market impact and the bid-ask spread. The optimization procedure tests 35 parameter combinations.[9]

Figure III-5: Simulation (in-sample) – cumulative dollar profit as a function of stop-loss factor *R* and profit-taking parameter *P*, as defined in Eq. (3) and (4)

Notes: Analysis is based on the development sample, which includes 393 Cramer events from January 1 to July 31, 2006, assuming US-$50,000 investment per event and 0.1% market impact and US-$20 broker commission per transaction, without reinvestment of profits. This graph is shown "as-is" from the strategy development process.



[9] R is set to values $[0.8\%, 1.0\%, .., 1.6\%]$ and P to values $[3.0\%, 3.5\%, .., 6.0\%]$.

Figure 5 displays the simulated absolute dollar profit for all combinations of R and P. The simulation results show that profitability is robust across all parameter settings. After consultation with the investors, we chose $P = 4.5\%$ and $R = 1.2\%$ as final parametrization. This is a tighter profit-target and a wider stop-loss than the profit maximum at $P = 5.5\%$ and $R = 1.0\%$ would suggest, due to investors' fears of failing to realize profits or closing otherwise profitable positions too early.

Figure III-6: Simulation (all datasets) – cumulative raw returns in and out of sample with parameters *R = 1.2%* and *P = 4.5%*

Notes: Analysis assumes 0.1% market impact and $20 broker commission per transaction for an investment of $50,000 per event, without reinvestment of profits. The chart displays simulated trading results for all 916 Cramer events, of which a total of 209 pass the entry filter as given in Eq. (2). The in-sample dataset includes the Cramer events from Jan 3, 2006 to July 31, 2006, "out-of-sample I" those from August 1, 2006 to October 17, 2006, and "out-of-sample II" comprises events from July 28, 2005 to December 23, 2005. This graph is shown "as-is" from the strategy development process.

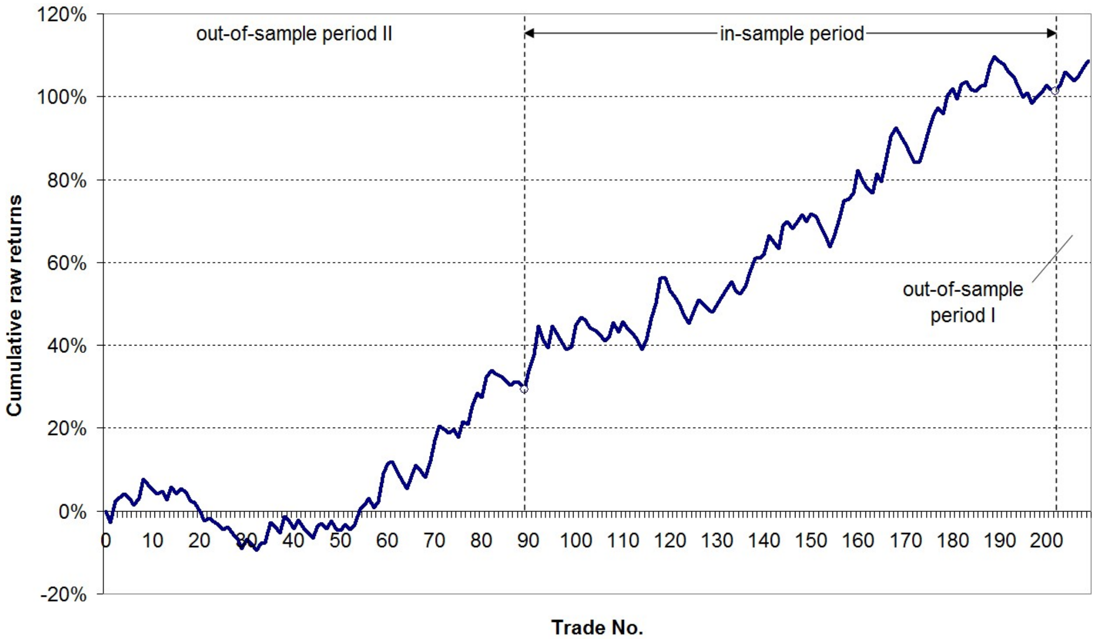


Figure 6 shows the final strategy's simulated account equity curve across all sample periods. During the in-sample period, assuming that every Cramer event can be traded using the full available capital, the strategy achieves a cumulative return of approximately 70.4%.[10] It retains its profitability in both out-of-sample periods. The Sharpe ratio is 3.05 and 2.00 in the in- and out-of-sample simulation, respectively, and

[10] For intraday short-sales, only 30% margin requirements apply, allowing up to three simultaneous positions.

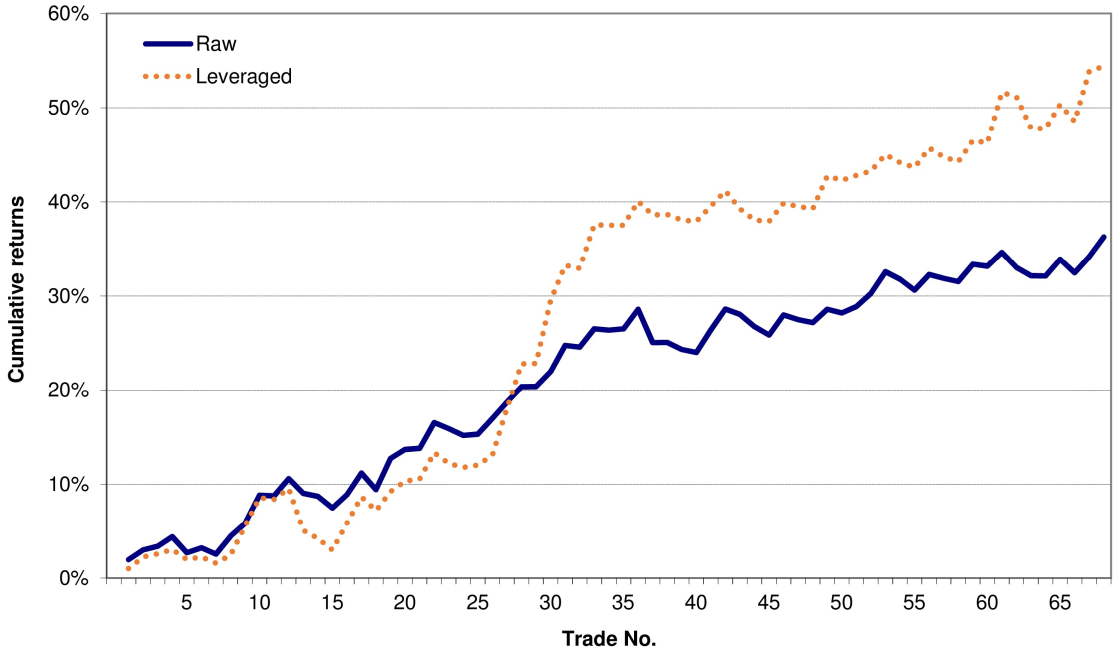
2.86 across all ex-ante test data. Table 2 in section 4 provides detailed statistics. These results motivated investors to provide capital for the trading experiment.

4. Results

We start by reporting the results of the actual-dollar trading test in subsection 4.1. In subsection 4.2 we investigate the results to understand the source of profits.

Figure III-7: Cumulative raw and leveraged returns of 68 actual-dollar trades from December 4, 2006 to July 27, 2007

Notes: Chart displays raw and leveraged returns for actual-dollar trades including transaction costs. Raw returns assume normalizing trades to a leverage of 1.0. Actual leverage during the trading experiment was a weighted average of 1.40. See Appendix for a detailed list of transactions.



4.1. Actual-Dollar Trading Results

The actual-dollar out-of-sample test started on December 4, 2006. The strategy was executed using an Advisor Account at Interactive Brokers, generating a total of 68 trades[11] until July 31, 2007 (see Appendix for a detailed list of transactions).

[11] In a joint decision with investors, the first eight trades were executed manually with smaller position sizes to observe potential market reaction to the arbitrage strategy. The 60 remaining trades were executed without

Table III-2: Trading strategy performance: trade-by-trade statistics and weekly benchmark comparison

Notes: This table compares actual-dollar trading results during the period from December 4, 2006 to July 27, 2007 with the backtest simulation (Panel A) and simple benchmark strategies (Panel B; using the Dow Jones Wilshire 5000 Index from Yahoo! Finance as a measure of total stock market performance). Panel A statistics are calculated on a per-trade basis, whereas Panel B reports weekly returns. Actual-dollar trades are reported including all transaction costs. Benchmark returns do not include transaction costs.

Panel A: Per-trade statistics - actual-dollar trades versus historic testing periods

	Actual-dollar trades (Dec. 4, 2006 - Jul. 27, 2007)		**Simulated trades (backtest results)** (Jul. 28, 2005 - Oct. 17, 2006)	
			In-sample	**Out-of-sample I+II**
Performance measure	raw returns*	leveraged**	raw returns*	raw returns*
Total trades	68	68	114	95
- thereof winning	39	39	59	45
- thereof losing	29	29	55	50
% winning trades	57.4%	57.4%	51.8%	47.4%
Cumulative return	**36.27%**	**54.42%**	**70.43%**	**38.24%**
Average trade profit	**0.53%**	**0.80%**	**0.62%**	**0.40%**
Average winner	1.56%	2.09%	2.79%	2.39%
Average loser	-0.84%	-0.93%	-1.71%	-1.38%
Win/loss ratio	1.84	2.24	1.63	1.73
Largest winner	3.30%	6.68%	6.76%	6.89%
Largest loser	-3.55%	-4.42%	-2.98%	-2.73%
Standard deviation	1.42%	2.08%	2.63%	2.22%
P-value	**0.001**	**0.001**	**0.006**	**0.039**
Sharpe Ratio* **	**3.30**	**3.20**	**3.05**	**2.00**
- simulation Sharpe Ratio***			ex-ante combined:	2.86
Avg. holding period (minutes)	34.3	34.3	n/a	n/a
Return on capital, per hour	0.93%	1.40%	n/a	n/a
Annualized Return	54.4%	81.6%	120.7%	70.6%

* equal-weighted trades with leverage normalized to 1.0

** taking into account actual leverage on investors' paid-in capital, weighted average leverage was 1.4

*** based on annualized weekly returns and standard deviation, assuming a risk free rate of 5.25% p.a.

Panel B: Comparison of actual-dollar trades to benchmarks (weekly returns from Dec. 4, 2006 - Jul. 30, 2007)

	Actual-dollar trades		**Benchmark returns**	
			DJ Wilshire 5000	**DJ Wilshire 5000**
Performance measure	raw returns*	leveraged**	(Long buy&hold)	(Short "sell&hold")°
Cumulative return	36.27%	54.42%	6.03%	-2.51%
Annualized return	54.40%	81.63%	9.04%	-3.77%
Annualized std. deviation	14.92%	23.89%	12.03%	12.03%
Sharpe Ratio (at 5.25% risk-free)	3.30	3.20	0.32	(0.75)
Maximum drawdown°°	-3.52%	-3.68%	-5.45%	-9.54%

° assuming the inverse of index buy&hold returns plus 5.25% per annum risk-free return on cash

°° defined as maximum "peak-to-through" loss of account equity

Figure 7 displays the realized cumulative raw returns for actual-dollar trades, normalizing invested capital to a leverage of 1.0 (blue line), and the weighted returns using the actual leverage (dotted orange line)[12]. The strategy has achieved a

manual intervention. The stop-loss limit was implemented using trailing stop orders offered by the broker (see order (2b) in Figure 4).

[12] Actual leverage ratios during trading ranged from 0.13 to 4.21 of paid-in capital, depending on liquidity of Cramer events and "pattern day-trading" limitations for U.S. stock trading with accounts under US-$25,000. We decided to limit position size to about 1.5% of the 20-day average dollar trading volume. Weighted-average leverage was 1.4 for all trades.

cumulative raw return of 36.3% with a Sharpe ratio of 3.30, and 54.4% leveraged return with Sharpe ratio of 3.20. Table 2 provides detailed performance statistics. Panel A shows a comparison of strategy performance between the actual-dollar trading period and the simulated results generated during ex-ante strategy development (see section 3.3). The results show a high significance of actual-dollar profitability, with p-values of .001. Performance is also highly persistent across periods. On an absolute basis, the average real trade of 0.53% is close to the optimized in-sample result of 0.62%. The same is true on a risk-adjusted basis as indicated by the Sharpe ratio. These results exclude the additional interest income which investors earned in the money market, since their trading capital was never used over night. The trade-based statistics show that the distribution of actual trade profits is doubly-positively skewed. The strategy incurs more winners (39) than losers (29) and the average profit of a winning trade (1.56%) is almost twice that of a losing trade (-0.84%). Comparing these statistics with the ex-ante simulation period, we find that both components improved. To put this in context, during simulation, the lack of intraday data forced us to make conservative assumptions whenever the daily $High_{t+1} - Low_{t+1}$ price range indicated that both the profit-target and the stop-loss level would have been triggered. Panel B compares the strategy's actual-dollar trades to the performance of simple benchmark strategies ("buy&hold" and "short&hold" the overall stock market measured by the Dow Jones Wilshire 5000 Index). The strategy widely outperforms in terms of absolute returns and even more so on a risk-adjusted basis. Figure 8 shows a graphical comparison of weekly strategy raw returns versus the two benchmark strategies.

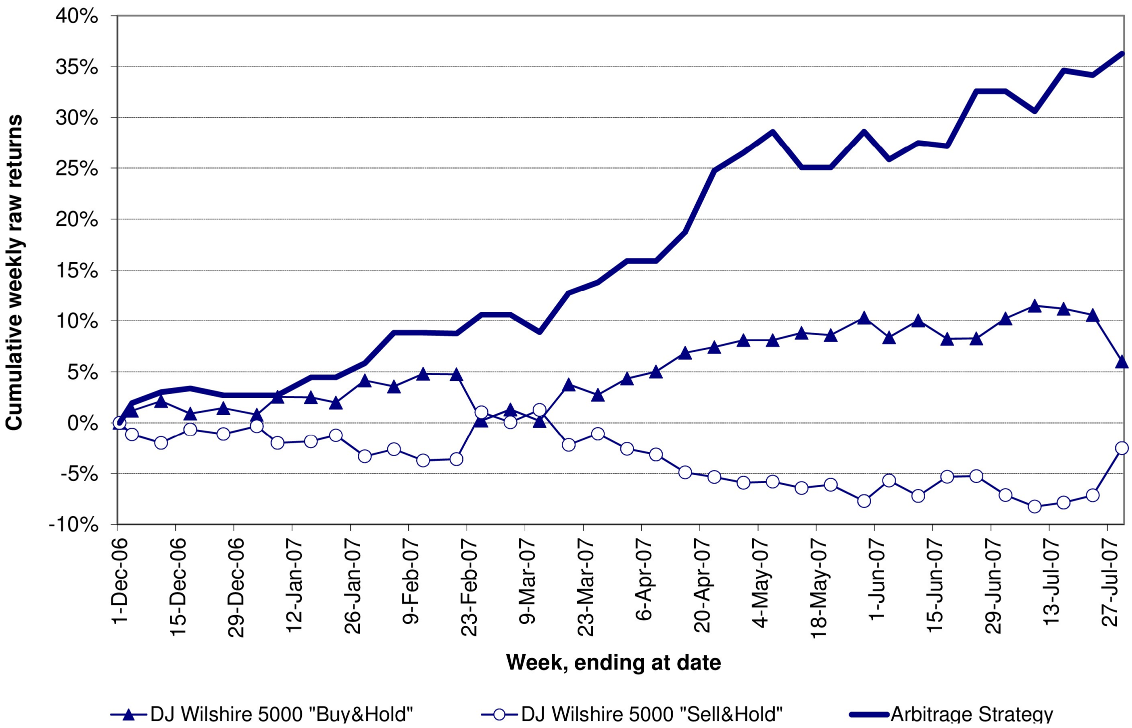
Across all performance measures, there is no doubt that the strategy's real-money out-of-sample test was successful.

4.2. Ex-Post Investigations and Implications for Market Efficiency

Actual-dollar trading results show that Cramer events represent an economically exploitable instance of return predictability in the market. But does this imply a true market inefficiency? We start by testing whether Cramer events allow earning *excess returns* after adjusting for known *risk factors* in the market efficiency literature. So far we have only looked at raw returns, adjusting only for the risks from the volatility of the strategy itself.

Figure III-8: Strategy benchmark comparison of weekly returns from December 1, 2006 to July 27, 2007

Notes: Chart displays cumulative weekly raw returns for actual-dollar transactions including transaction costs, normalizing trades to a leverage of 1.0. Actual leverage on investors' paid-in capital during the trading experiment was 1.40 on average (see Appendix for detailed data on individual transactions). Benchmark strategies are defined as holding long or short the Dow Jones Wilshire 5000 Index from December 1, 2006 to July 27, 2007, whereby the short benchmark strategy is assumed to earn a 5.25% p.a. risk-free rate on the cash received from the short-sale.



To investigate this, we follow Fama & French (1992) and assess if the strategy's returns might be explained by known risk factors[13]: (1) the "Market" risk factor, (2) the "Size" factor capturing the risk of investing in small firms, (3) the "Value" factor capturing the risk of investing in firms with low book-to-market ratios, (4) the "Momentum" factor (Jeegadeesh & Titman, 1993) capturing the risks of investing in firms with large relative past long-term returns, and finally (5) the "Short-term Reversal" factor capturing the risk of investing in firms with the lowest past short-term returns, which has been found to proxy for *liquidity risk* (see e.g. Pástor & Stambaugh, 2003; Amini et al., 2013).

To reduce the impact of the fact that the strategy does not trade every day, we sum up unleveraged trade returns to weekly returns and regress these against contemporaneous returns of Market, Size, Value, Momentum and Short-term Reversal factors. The specification is:

$$\begin{aligned} r_t &= \alpha + \beta * Market_t + \chi * Size_t + \delta * Value_t \\ &+ \gamma * Momentum_t + \eta * Reversal_t + \varepsilon_t, \end{aligned} \tag{5}$$

whereby r_t is the strategy's weekly return. As a cross-validation of our results, we refer to Bolster & Trahan (2009) who investigate the style effects present in Jim Cramer's picks. The authors find that Cramer events tend to be smaller value stocks with positive momentum. Hence, since we are *short* these stocks, if only for a very short time, we should get *negative* coefficients on the Market, Size, Value, and

[13] As defined in the Kenneth French Data Library: The *"Market"* risk factor is the value-weighted return of all NYSE, AMEX and NASDAQ stocks minus the risk free rate. The *"Size"* factor is going long firms below the median market capitalization versus short the firms above. The *"Value"* factor is going long a 50:50 mix of value-weighted small and large firm portfolios below the 30th percentile of book-to-market ratios (book equity divided by market capitalization), and vice-versa short firms above the 70th percentile. The *"Momentum"* factor is going long a 50:50 mix of value-weighted small and large firm portfolios above the 70th percentile of prior returns measured from day -250 to -21, and vice-versa short firms below the 30th percentile. The *"Short-term Reversal"* factor is defined as going long a 50:50 mix of value-weighted small and large firm portfolios up to the 30th percentile of prior returns measured from day -20 to -1, and vice-versa short firms above the 70th percentile.

Momentum factors. In contrast, the strategy's shorting against price spikes after Cramer events should be correlated to the short-side of the Reversal factor portfolio; hence, we should find a *positive* coefficient. Indeed, the regression results in Table 3 confirm these expectations.

Table III-3: Regression of strategy returns against common risk factors

*Notes: This table shows regression results and descriptive statistics for weekly strategy returns against common risk factors, using the regression specification in Eq. (5). Return data for Market, Size, Value, Momentum and Return Reversal risk premia are taken from Kenneth French's data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). Panel A covers the 35 weeks in the actual-dollar trading period from December 4, 2006 to July 30, 2007. Panel B shows results for the whole period from July 28, 2005 to July 30, 2007, combining simulated and actual-dollar results and using only the 79 weeks in which the trading strategy performed at least one trade. Significant coefficients are highlighted in **bold**.*

Panel A: Actual trading results - all weeks including inactive

(Dec. 4, 2006 - Jul. 30, 2007, N=35 weekly returns)

	Regression results			Sample statistics	
	Coeff.	t-stat	P(>|t|)	Average	Std.dev.
Strategy return				1.04	2.07
Unexplained return ("alpha")	**1.13**	**3.01**	**0.010****	1.13	1.92
Risk Factors					
Market (minus Risk-free)	-0.00	-0.01	0.995	0.11	1.74
Size	-0.08	-0.18	0.861	-0.11	0.84
Value	-0.75	-0.99	0.331	-0.07	0.55
Momentum	-0.83	-1.37	0.180	0.16	0.67
Short-term Reversal	0.51	0.81	0.422	-0.03	0.71
R^2	13%				

Panel B: Simulation and actual results combined - active trading weeks only

(Jul. 28, 2005 - Jul. 30, 2007, N=79 weekly returns)

	Regression results			Sample statistics	
	Coeff.	t-stat	P(>|t|)	Average	Std.dev.
Strategy return				1.83	3.78
Unexplained return ("alpha")	**1.98**	**4.78**	**<0.0001*****	1.98	3.48
Risk Factors					
Market (minus Risk-free)	-0.45	-1.30	0.197	0.16	1.44
Size	-0.16	-0.31	0.761	-0.01	0.90
Value	-0.22	-0.32	0.751	0.06	0.64
Momentum	-0.41	-0.91	0.364	0.05	1.14
Short-term Reversal	**1.29**	**2.64**	**0.010***	-0.03	0.84
R^2	15%				

Significance Levels: *** < 0.0001, ** < 0.01, * < 0.05

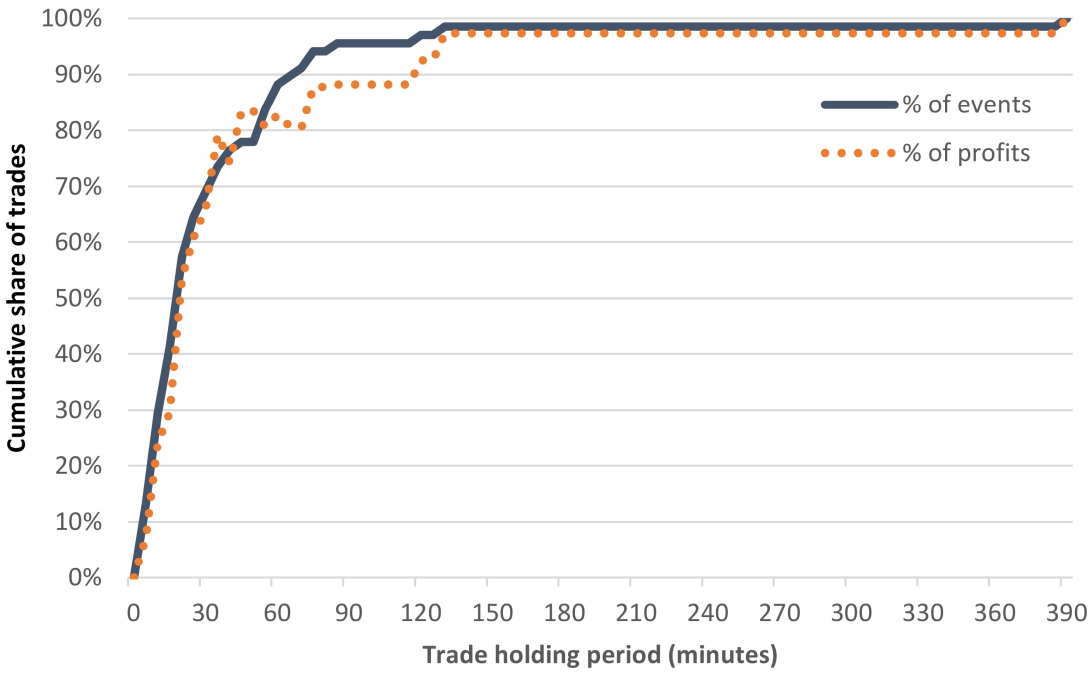
Panel A shows the results for a sample covering only the 35 weeks of live trading, which is likely too short to obtain significant parameter estimates. Except for the regression intercept (unexplained weekly strategy return), none of the coefficients is even barely significant. Therefore we extend our sample to 79 weeks in Panel B to include the in-sample simulation period as well, while removing the noise from those weeks in which the strategy did not trade. Now the Reversal factor becomes a significant predictor of the strategy's return. However, taking into account the descriptive statistics for the variables, it becomes clear that all this explains is small *variations* of weekly strategy returns around their significantly positive mean of 1.83%. The strategy's unexplained return (its "Alpha") is by far the most significant predictor of strategy returns!

Based on the standard risk factor regression, we cannot refute that the strategy earned significant excess returns, which points to market inefficiency. On the other hand, it is well possible that mapping the returns of an intraday strategy to weekly returns and then regressing these returns against contemporaneous factor returns is too crude an instrument to explain the components of intraday profits. Unfortunately, to our knowledge, the market efficiency literature does not provide a model for intraday expected returns.

How the trading rule navigated the intraday price paths of Cramer events should provide clues whether the strategy was merely paid for taking risks, or whether it earned excess returns from market timing. The detailed list of transactions (see Appendix) reports the holding period for each trade, in minutes. Surprisingly, we do not find a relatively uniform distribution of holding periods with a mix of longer and shorter trades. To illustrate the data, Figure 9 depicts the cumulative sums of trade count and profit for trades sorted from shortest to longest holding periods. There is a clear concentration of trade exits in the first *minutes* of the trading day. More than half of trades are closed within the first 20 minutes, about two-thirds during the first half-hour. The average holding period is only 34 minutes. No trades except one make it to the market close. Rather, there is a substantial bias that either the strategy's profit-taking order or the trailing stop-loss order is triggered early in the day.

Figure III-9: Cumulative trades and actual-dollar profits by holding period

Notes: Chart displays cumulative share of trade count and trade profits, sorting the 68 actual-dollar trades by holding period.



At the same time, these shorter trades are no less profitable than their longer counterparts. Moreover, most of the trades exited via trailing stop were exited at small losses or even small profits. Due to the logic of the trailing stop, which follows prices down at a distance, this implies that many of these trades actually must have been profitable trades which reversed *upwards* to trigger the trailing stop level (see order (2b) in Figure 4 to illustrate this). In other words, after the initial down-move to correct the initial mispricing from Cramer events, there should be a systematic tendency of stocks to rise again intraday. In short, a *"reversal-of-the-reversal"* pattern. To test for the existence of this pattern, we obtain intraday price data for a second sample of 105 Cramer events with market capitalization of up to US-$10 billion and Cramer price impacts (CPI) of at least 1% at market opening, covering trading days from November 16, 2006 to May 29, 2007[14], along with intraday prices of the SPDR S&P 500 ETF[15]. We perform separate regressions of the event returns in the first and second half-hours of the trading day, respectively. For the *first half-hour*, we use the following specification:

$$r_{i,t[9:30-9:59]} = \alpha + \omega * Overnight_{i,t} + \beta * Market_{i,t[9:30-9:59]} + \varepsilon_{i,t}, \qquad (6)$$

whereby $r_{i,t}$ is a Cramer event's return in the first half-hour of the day, $Overnight_{i,t}$ is the close-to-open return $O_{i,t+1}/C_{i,t} - 1$ and $Market_{i,t[9:30-9:59]}$ is the contemporaneous market return in the first half-hour using the prices of the SPDR S&P 500 ETF as a proxy. For the *second half-hour*, we shift all terms except overnight return forward by half an hour, and add the first half-hour stock return as a variable to test for intraday reversal:

$$\begin{aligned} r_{i,t[10:00-10:30]} &= \alpha + \rho * r_{i,t[9:30-9:59]} + \omega * Overnight_{i,t} \\ &+ \beta * Market_{i,t[10:00-10:30]} + \varepsilon_{i,t}, \end{aligned} \qquad (7)$$

Table 4 shows that there is indeed a significant tendency for price declines of the first half-hour to be reversed upwards in the second half-hour. As expected, in the first half-hour the initial down-move partially reverses the overnight price-spike. The market factor has a significant influence as well with an intraday "beta" of about 2. In the second half-hour, however, instead of continuing to correct the mispricing, there is a significant tendency to *reverse upwards* part of the first half-hour down-move. The

[14] Provided by Bloomberg. The sample contains 33 of the actual trades.

[15] Symbol: SPY, provided by Lenz & Partner AG.

coefficient of -.22 is highly significant at the .0001 level. The size of the initial mispricing ($Overnight_{i,t}$) is also a significant predictor, suggesting that the initial speculative interest from retail investors (see Engelberg et al., 2012) might also be behind this reversal-of-the-reversal phenomenon.

Table III-4: Regressions of first and second half-hour event returns

Notes: This table shows regression results and descriptive statistics for regressions of first and second half-hour returns on event and market return variables. The intraday data comprises prices by Bloomberg for a sample of 105 Cramer Events with overnight returns of at least 1% and market capitalization of <= US-$10 billion, along with contemporaneous market return using the SPDR S&P 500 ETF as a proxy. The Cramer Events cover a period with trading days from November 16, 2006 to May 29, 2007. Panel A regresses stock returns in the 1st half-hour of the trading day against the size of the overnight jump and the market return in the 1st half-hour (see Eq. (6)). Panel B regresses stock returns in the 2nd half-hour of the trading day against stock returns in the 1st half-hour, the size of the overnight jump and the market return in the 2nd half-hour (see Eq. (7)).

Panel A: Regression of stock returns in the *first half-hour* of the day (N = 105)

	Regression results			**Sample statistics**	
	Coeff.	t-stat	P(>|t|)	Average	Std.dev.
Regressand: *1st half-hour* return				-0.15	1.09
Intercept	-0.10	-0.48	0.630		
Regressors					
Overnight return *O_t+1-C_t0*	-0.09	-2.51	0.014*	2.19	2.99
Market 1st half-hour return	2.01	2.15	0.034*	0.00	0.13
R^2	10%				

Panel B: Regression of stock returns in the *second half-hour* of the day (N = 105)

	Regression results			**Sample statistics**	
	Coeff.	t-stat	P(>|t|)	Average	Std.dev.
Regressand: *2nd half-hour* return				0.05	0.73
Intercept	-0.18	-1.63	0.106		
Regressors					
Stock 1st half-hour return	-0.22	-4.33	<0.0001***	-0.15	1.09
Overnight return *O_t+1-C_t0*	0.05	2.55	0.012*	2.19	2.99
Market 2nd half-hour return	1.17	3.15	0.002**	-0.01	0.21
R^2	28%				

Significance Levels: *** < 0.0001, ** < 0.01, * < 0.05

Did the strategy manage to earn *excess returns* from this intraday pattern? We propose a simple benchmark of what the fair compensation for a liquidity-providing arbitrageur should be. Studies on stock recommendations[16] and Cramer events[17] show that most if not all abnormal post-recommendation returns subside after a few days, weeks or months. Since our strategy holds the stocks only during the first day ($t+1$), a prudent assumption on the *expected* return for the *average* arbitrageur on the trading day would be the expected down-move from $Open_{t+1}$ to $Close_{t+1}$.[18] Using this simple benchmark, the expected profit amounts to about 33.5% earned with a holding period of six and a half hours for the 68 observed trades. The strategy however, earned 36.3% with an average holding period of only 34 minutes. Therefore, not only did the strategy manage to capture a slightly higher-than-expected profit, it also put its capital at risk for less than one-tenth of the time. This suggests that next to a fair compensation for liquidity provision, the strategy also earned excess returns from timing the market better intraday than the average market participant.

5. Discussion and Conclusions

This paper conducts an actual-dollar trading experiment as a data-snooping bias-free and realistic out-of-sample test of an ex-ante-specified trading rule. We exploit the correction of mispricings caused by retail investors buying stocks after recommendations by Jim Cramer in the daily evening TV show "Mad Money" on CNBC. Investigating the source of these profits suggests that common risk factors cannot refute that the strategy earns excess returns, although some exposure to short-term Reversal indicates that it earns a premium for liquidity provision. Looking

[16] see e.g. Kerl & Walter (2007), Busse & Green (2002), Liang (1999), Metcalf & Malkiel (1994), Wright (1994), and Barber & Loeffler (1993)

[17] see e.g. Engelberg et al. (2012), Bolster et al. (2012), Hobbs et al. (2012), Chen et al. (2011), Keasler & McNeil (2010), Lim & Rosario (2010), Bolster & Trahan (2009), Karniouchina et al. (2009), and Neumann & Kenny (2007)

[18] This definition is conceptually comparable with the statistical arbitrageurs in Seasholes & Wu (2007), who accumulate shares on day t0 to sell on day t+1. For Cramer events, t0 is before the recommendation, hence the Open on t+1 is the earliest possible entry point.

intraday confirms that the strategy also benefited from market timing, by navigating a "reversal-of-the-reversal" pattern early in the day better than a benchmark average arbitrageur.

These findings contribute to several literatures in financial economics. First, in the *market efficiency* literature, this study appears to be the first to use the stock market as a laboratory to test the validity of a trading rule. This provides a realistic out-of-sample test without data-snooping bias. Furthermore, the high share of unexplained profits calls for further research on intraday expected returns. Second, in the literature on *recommendation effects*, this study confirms that the abnormal returns found in prior studies on Cramer events are actually exploitable. The intraday "reversal-of-the-reversal" adds to previous results showing only a monotonous decline of the mispricing. In fact, we find that the *mispricing re-widens* intraday. This implies the need for arbitrageurs to time the market to maximize profits from the mispricing. To the third and final body of literature on *limits to arbitrage*, these findings provide confirmatory evidence for theoretical models. The intraday re-widening of mispricings supports Abreu & Brunnermeier (2002) who suggest that arbitrageurs might fail to correct a mispricing because they fail to synchronize their trades. In their model, no individual arbitrageur has enough risk capacity to counter the demand shock following a stock recommendation, and therefore an arbitrageur has to rely on the actions of other investors moving prices in the desired direction. This coordination problem leads to strategic uncertainty as to whether the arbitrage will be successful or not, causing arbitrageurs to delay their transactions as they attempt to time the market. Moreover, knowing about the transitory intraday up-trend makes it rational to speculate in the direction of the mispricing (i.e. *contrary* to fundamentals), thus increasing the mispricing further. This supports the hypothesis of "destabilizing speculation" by De Long et al. (1990). Taken together, part of the high returns of the trading experiment might well be explained by suboptimal strategic behavior of less sophisticated arbitrageurs.

Finally, do the excess returns also constitute a significant market inefficiency? In a conference presentation on the neoclassical response to market anomalies, Stephen Ross proposes the following test for a new anomaly (Ross, 2006):

(1) Is it true? (i.e. are returns statistically significant and not explained by other factors?)

(2) How damaging is it to the neoclassical theory of finance? (i.e. is it economically large and refuting existing models of asset prices?)

According to Ross, we should only be concerned by phenomena that score high in both dimensions. On the first dimension, this paper shows statistically significant excess returns that cannot be explained by risk factors. On the second dimension, however, the excess returns and intraday patterns found after Cramer events do not constitute an economically large deviation from efficient prices – at least not on a *stand-alone* basis. To illustrate, consider that the median daily liquidity of traded stocks is only about US-$25 to 50 million per day(see Table I). Even if we could trade five percent of a stock's liquidity without market impact destroying profitability, the strategy would still only scale to an invested capital of US-$2.5 million per trade. If we were content with 5% annual returns, maybe we could deliver these returns for US-$10 to 20 million of capital. Still, this is orders of magnitudes smaller than the break-even sizes of well-known anomalies such as Value or Momentum, which reach hundreds of billions (Frazzini et al., 2012; Asness et al., 2014). For a large institutional investor, Cramer events would be a tiny profit opportunity.

Further investigations of recommendations and other attention-based events such as large price changes are necessary to confirm and generalize our results. For the example of Cramer events, we learn that the "business" of correcting mispricings is profitable, risky and competitive.

Acknowledgements

I thank the investors, whose commitment of over EUR 200.000 of trading capital made this trading experiment possible. I thank professors Günter Bamberg and Yarema Okhrin for helpful discussions and suggestions. I thank Emily Niemann and Ronnie Sadka for helpful comments.

Appendix: Detailed Account of Actual-Dollar Trades

Trade No.	Symbol	Recommendation date	Trade date	Position size	Short-sale price	Investment	Profit target	Stop distance	Trailing trigger**	Exit Type	Market impact costs	Realized exit price	Gross profit	Broker fees	Net profit		Holding period	Paid-in investor capital	Implied leverage	Return on capital
				(shares)	($)	($)	($)	($)	($)	(***)	%	($)	($)	($)	($)	(%)	(mins)	($)		(%)
1	AUY	12/01/06	12/04/06	500	13.30	6,650	12.80	0.11	13.41	man*	0.00%	13.030	135	5	130	1.95%	60	13,000	0.51x	1.00%
2	DLB	12/07/06	12/08/06	500	30.63	15,315	29.31	0.33	30.96	man*	0.00%	30.298	166	5	161	1.05%	35	13,000	1.18x	1.24%
3	GSIC	12/14/06	12/15/06	1,900	18.79	35,701	17.82	0.34	19.13	man*	0.00%	18.711	151	19	132	0.37%	30	39,000	0.92x	0.34%
4	GMKT	12/18/06	12/19/06	800	21.68	17,344	20.75	0.32	22.00	man*	0.00%	21.446	187	8	179	1.03%	10	39,000	0.44x	0.46%
5	ELOS	12/20/06	12/21/06	1,100	27.70	30,470	27.20	0.46	28.16	SL*	0.00%	28.160	-506	11	-517	-1.70%	5	48,750	0.63x	-1.06%
6	ECOL	01/09/07	01/10/07	1,300	19.17	24,921	18.50	0.41	19.58	man*	0.00%	19.062	140	13	127	0.51%	5	48,750	0.51x	0.26%
7	PETM	01/10/07	01/11/07	1,500	30.95	46,425	30.50	0.23	31.18	man*	0.00%	31.149	-298	15	-313	-0.67%	15	48,750	0.95x	-0.64%
8	MRVC	01/15/07	01/16/07	5,100	4.17	21,267	3.94	0.09	4.26	man*	0.00%	4.080	458	51	407	1.91%	25	48,750	0.44x	0.83%
9	NTLI	01/23/07	01/24/07	4,000	27.22	108,880	26.83	0.17	27.00	PT	0.00%	26.830	1,560	40	1,520	1.40%	10	48,750	2.23x	3.12%
10	WSII	01/30/07	01/31/07	5,300	11.75	62,275	11.39	0.26	11.65	PT	0.00%	11.390	1,908	53	1,855	2.98%	75	61,750	1.01x	3.00%
11	LMS	02/13/07	02/14/07	4,200	29.75	124,950	28.87	0.64	29.76	TS	0.01%	29.763	-55	42	-97	-0.08%	70	61,750	2.02x	-0.16%
12	CHTR	02/22/07	02/23/07	11,700	3.26	38,142	3.19	0.05	3.24	PT	0.00%	3.190	819	117	702	1.84%	2	61,750	0.62x	1.14%
13	PRAA	03/05/07	03/06/07	4,000	43.84	175,360	42.67	0.73	44.12	TS	0.89%	44.513	-2,692	40	-2,732	-1.56%	8	61,750	2.84x	-4.42%
14	BGC	03/06/07	03/07/07	3,500	50.75	177,625	49.78	0.48	50.83	TS	0.13%	50.896	-511	35	-546	-0.31%	60	61,750	2.88x	-0.88%
15	FSLR	03/07/07	03/08/07	1,200	49.73	59,676	48.84	0.50	50.16	TS	0.38%	50.350	-744	12	-756	-1.27%	2	61,750	0.97x	-1.22%
16	ESRX	03/07/07	03/08/07	1,600	77.15	123,440	76.04	0.41	76.45	PT	0.00%	76.040	1,776	16	1,760	1.43%	20	61,750	2.00x	2.85%
17	NCMI	03/16/07	03/19/07	2,700	27.80	75,060	27.15	0.45	27.60	PT	0.00%	27.150	1,755	27	1,728	2.30%	20	61,750	1.22x	2.80%
18	JSDA	03/16/07	03/19/07	2,500	20.35	50,875	19.76	0.43	20.64	TS	0.29%	20.70	-875	25	-900	-1.77%	20	61,750	0.82x	-1.46%
19	NSTK	03/16/07	03/19/07	3,300	11.50	37,950	11.11	0.29	11.40	PT	0.00%	11.110	1,287	33	1,254	3.30%	45	61,750	0.61x	2.03%
20	HAYN	03/21/07	03/22/07	1,000	71.65	71,650	69.83	1.26	71.51	EOD	0.00%	70.949	701	10	691	0.96%	390	61,750	1.16x	1.12%
21	BRCD	03/23/07	03/26/07	10,600	10.40	110,240	10.23	0.10	10.35	TS	0.27%	10.378	233	106	127	0.12%	20	62,833	1.75x	0.20%
22	CPHD	03/26/07	03/27/07	5,500	11.70	64,350	11.37	0.21	11.58	PT	0.00%	11.370	1,815	55	1,760	2.74%	20	62,833	1.02x	2.80%
23	PDLI	03/26/07	03/27/07	5,300	20.93	110,929	20.52	0.21	20.96	TS	0.45%	21.054	-657	53	-710	-0.64%	55	62,833	1.77x	-1.13%
24	INCY	04/09/07	04/10/07	5,000	7.52	37,600	7.32	0.15	7.65	TS	0.00%	7.565	-225	50	-275	-0.73%	65	62,833	0.60x	-0.44%
25	ROSE	04/09/07	04/10/07	6,000	22.82	136,920	20.28	0.30	22.77	TS	0.04%	22.778	252	60	192	0.14%	30	72,930	1.88x	0.26%
26	CHTR	04/10/07	04/11/07	15,000	2.98	44,700	2.92	0.05	2.97	PT	0.00%	2.920	900	150	750	1.68%	10	72,930	0.61x	1.03%
27	NLY	04/12/07	04/13/07	13,000	16.19	210,470	15.90	0.17	16.07	PT	0.00%	15.900	3,770	130	3,640	1.73%	130	72,930	2.89x	4.99%
28	FWLT	04/16/07	04/17/07	3,000	70.82	212,460	69.66	0.61	70.27	PT	0.00%	69.660	3,480	30	3,450	1.62%	10	72,930	2.91x	4.73%
29	SGR	04/16/07	04/17/07	3,000	32.15	96,450	31.53	0.38	32.06	TS	0.22%	32.130	60	30	30	0.03%	25	72,930	1.32x	0.04%
30	MNST	04/17/07	04/18/07	10,905	44.82	488,773	44.10	0.35	44.45	PT	0.00%	44.100	7,863	109	7,753	1.59%	120	116,100	4.21x	6.68%
31	FTEK	04/20/07	04/23/07	7,000	23.67	165,690	23.00	0.49	23.49	PT	0.00%	23.000	4,690	70	4,620	2.79%	10	122,850	1.35x	3.76%
32	DISCA	04/23/07	04/24/07	11,156	21.33	237,957	21.00	0.17	21.32	TS	0.17%	21.357	-301	112	-413	-0.17%	15	122,850	1.94x	-0.34%
33	CAKE	04/25/07	04/26/07	10,000	29.49	294,900	28.91	0.36	29.27	PT	0.00%	28.910	5,800	100	5,700	1.93%	35	122,850	2.40x	4.64%
34	KGC	04/26/07	04/27/07	3,400	13.55	46,070	13.33	0.12	13.56	TS	0.00%	13.560	-34	34	-68	-0.15%	55	122,850	0.38x	-0.06%
35	PENN	04/27/07	04/30/07	227	48.58	11,028	47.74	0.46	48.49	TS	0.02%	48.500	18	2	16	0.14%	85	122,850	0.09x	0.01%
36	GSS	05/04/07	05/07/07	31,000	4.80	148,800	4.69	0.08	4.77	PT	0.00%	4.690	3,410	310	3,100	2.08%	15	122,850	1.21x	2.52%

Trade No.	Symbol	Recommendation date	Trade date	Position size	Short-sale price	Investment	Profit target	Stop distance	Trailing trigger**	Exit Type	Market impact costs	Realized exit price	Gross profit	Broker fees	Net profit	Net profit	Holding period	Paid-in investor capital	Implied leverage	Return on capital
				(shares)	($)	($)	($)	($)	($)	(***)	%	($)	($)	($)	($)	(%)	(mins)	($)		(%)
37	PAL	05/08/07	05/09/07	4,500	**11.00**	49,500	10.69	0.27	11.27	TS	0.98%	**11.380**	-1,710	45	-1,755	**-3.55%**	10	122,850	0.40x	**-1.43%**
38	CBI	05/09/07	05/10/07	5,500	**39.76**	218,680	39.06	0.27	39.66	TS	0.20%	**39.740**	110	55	55	**0.03%**	2	122,850	1.78x	**0.04%**
39	TSS	05/21/07	05/22/07	3,000	**33.60**	100,800	33.07	0.25	33.84	TS	0.00%	**33.840**	-720	30	-750	**-0.74%**	15	122,850	0.82x	**-0.61%**
40	MMR	05/22/07	05/23/07	3,140	**15.15**	47,571	14.69	0.34	15.19	TS	0.00%	**15.190**	-126	31	-157	**-0.33%**	75	122,850	0.39x	**-0.13%**
41	GLS	05/23/07	05/24/07	2,900	**28.45**	82,511	27.76	0.48	28.24	PT	0.00%	**27.760**	2,007	29	1,978	**2.40%**	20	122,850	0.67x	**1.61%**
42	TTEK	05/23/07	05/24/07	4,600	**22.48**	103,408	21.97	0.34	22.31	PT	0.00%	**21.970**	2,346	46	2,300	**2.22%**	20	141,750	0.73x	**1.62%**
43	CLF	05/30/07	05/31/07	7,700	**84.27**	648,879	82.75	0.86	84.61	TS	0.17%	**84.750**	-3,696	77	-3,773	**-0.58%**	20	196,701	3.30x	**-1.92%**
44	FCSX	05/31/07	06/01/07	4,000	**44.50**	178,000	43.36	0.80	44.63	TS	0.99%	**45.070**	-2,280	40	-2,320	**-1.30%**	15	196,701	0.90x	**-1.18%**
45	CCIX	05/31/07	06/01/07	1,000	**28.48**	28,480	27.64	0.61	28.46	TS	0.91%	**28.720**	-240	10	-250	**-0.88%**	60	196,701	0.14x	**-0.13%**
46	B	06/05/07	06/06/07	5,600	**31.79**	178,024	31.10	0.44	31.54	PT	0.00%	**31.100**	3,864	56	3,808	**2.14%**	10	196,701	0.91x	**1.94%**
47	ACAD	06/05/07	06/06/07	10,300	**14.12**	145,436	13.59	0.31	13.97	TS	1.50%	**14.180**	-618	103	-721	**-0.50%**	10	196,701	0.74x	**-0.37%**
48	CVTX	06/15/07	06/18/07	12,400	**11.92**	147,808	11.61	0.23	11.87	TS	0.67%	**11.950**	-372	124	-496	**-0.34%**	40	196,701	0.75x	**-0.25%**
49	BTU	06/18/07	06/19/07	9,400	**51.39**	483,066	50.65	0.28	50.93	PT	0.00%	**50.650**	6,956	94	6,862	**1.42%**	30	196,701	2.46x	**3.49%**
50	DYN	06/20/07	06/21/07	25,400	**9.31**	236,474	9.17	0.06	9.33	TS	0.05%	**9.335**	-635	254	-889	**-0.38%**	20	196,701	1.20x	**-0.45%**
51	BAM	06/20/07	06/21/07	3,800	**39.43**	149,834	38.87	0.21	39.11	TS	0.10%	**39.150**	1,064	38	1,026	**0.68%**	15	196,701	0.76x	**0.52%**
52	VGR	06/20/07	06/21/07	3,000	**21.42**	64,260	20.92	0.34	21.26	PT/TS	0.00%	**21.116**	912	30	882	**1.37%**	55	196,701	0.33x	**0.45%**
53	OMTR	06/22/07	06/25/07	6,500	**22.85**	148,525	22.31	0.38	22.69	PT	0.00%	**22.310**	3,510	65	3,445	**2.32%**	5	196,701	0.76x	**1.75%**
54	TX	07/06/07	07/09/07	6,500	**32.41**	210,665	31.90	0.24	32.55	TS	0.33%	**32.658**	-1,609	65	-1,674	**-0.79%**	10	196,701	1.07x	**-0.85%**
55	HNR	07/06/07	07/09/07	6,500	**12.85**	83,525	12.47	0.28	12.98	TS	0.08%	**12.990**	-910	65	-975	**-1.17%**	20	196,701	0.42x	**-0.50%**
56	TCK	07/10/07	07/11/07	5,200	**46.44**	241,488	45.77	0.25	45.85	PT	0.00%	**45.660**	4,056	52	4,004	**1.66%**	0.3	196,701	1.23x	**2.04%**
57	EMC	07/11/07	07/12/07	23,000	**19.50**	448,500	19.22	0.11	19.51	TS	0.31%	**19.570**	-1,610	230	-1,840	**-0.41%**	5	196,701	2.28x	**-0.94%**
58	APD	07/11/07	07/12/07	3,500	**84.20**	294,700	82.99	0.25	84.27	TS	0.25%	**84.480**	-980	35	-1,015	**-0.34%**	5	196,701	1.50x	**-0.52%**
59	GPRO	07/12/07	07/13/07	3,900	**61.93**	241,527	60.77	0.65	61.42	PT	0.00%	**60.770**	4,524	39	4,485	**1.86%**	10	196,701	1.23x	**2.28%**
60	APA	07/12/07	07/13/07	3,000	**86.70**	260,100	85.45	0.46	86.77	TS	0.12%	**86.870**	-510	30	-540	**-0.21%**	25	196,701	1.32x	**-0.27%**
61	TEX	07/13/07	07/16/07	8,200	**95.56**	783,592	94.17	0.53	94.70	PT	0.00%	**94.170**	11,398	82	11,316	**1.44%**	20	210,201	3.73x	**5.38%**
62	WGOV	07/16/07	07/17/07	1,100	**61.40**	67,540	60.15	0.78	62.18	TS	0.30%	**62.366**	-1,063	11	-1,074	**-1.59%**	55	210,201	0.32x	**-0.51%**
63	ASD	07/16/07	07/17/07	12,500	**62.75**	784,375	61.85	0.34	62.09	TS	1.95%	**63.300**	-6,875	125	-7,000	**-0.89%**	25	210,201	3.73x	**-3.33%**
64	SPR	07/17/07	07/18/07	10,500	**40.56**	425,838	39.90	0.32	40.33	TS	0.55%	**40.550**	63	105	-42	**-0.01%**	12	210,201	2.03x	**-0.02%**
65	NCX	07/19/07	07/20/07	7,200	**41.59**	299,448	40.84	0.43	41.27	PT	0.00%	**40.840**	5,400	72	5,328	**1.78%**	25	210,201	1.42x	**2.53%**
66	TSRA	07/20/07	07/23/07	5,900	**43.73**	258,007	42.84	0.56	44.12	TS	0.52%	**44.350**	-3,658	59	-3,717	**-1.44%**	40	210,201	1.23x	**-1.77%**
67	CY	07/20/07	07/23/07	25,300	**25.78**	652,234	25.33	0.25	25.58	PT	0.00%	**25.330**	11,385	253	11,132	**1.71%**	35	210,201	3.10x	**5.30%**
68	SCMR	07/26/07	07/27/07	13,600	**4.29**	58,344	4.19	0.07	4.26	PT	0.00%	**4.190**	1,360	136	1,224	**2.10%**	15	210,201	0.28x	**0.58%**
Total													**67,779**	**4,464**	**63,315**	36.27%	**2,331**			54.42%

* manual profit taking, fixed stop-loss level at Short-sale price + Stop distan

** stop-loss limit price at the time when the trailing stop-loss was triggered. The activation of the stop-loss order immediately sends a "buy at market" order to close the short position

*** PT = profit target, TS = trailing stop-loss (stop-buy order), SL = fixed stop-loss (stop-buy order), EOD = exit at market close, man = manual profit taking via market and limit buy orders

References (Chapter III)

Abreu, D., Brunnermeier, M., 2002. Synchronization risk and delayed arbitrage. Journal of Financial Economics 66, 341–360.

Amini, S., Gebka, B., Hudson, R., Keasey, K., 2013. A review of the international literature on the short term predictability of stock prices conditional on large prior price changes: Microstructure, behavioral and risk related explanations. International Review of Financial Analysis 26, 1–17.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T., 2014. Fact, fiction and momentum investing. Journal of Portfolio Management 40, 75–92.

Avramov, D., Chordia, T., 2006. Predicting stock returns. Journal of Financial Economics 82, 387–415.

Barber, B., Loeffler, D., 1993. The "Dartboard" column: Second-hand information and price pressure. The Journal of Financial and Quantitative Analysis 28, 273–284.

Barber, B., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Review of Financial Studies 21, 785–818.

Bolster, P., Trahan, E., Venkateswaran, A., 2012. How mad is Mad Money? Jim Cramer as a stock picker and portfolio manager. Journal of Investing 21, 27–39.

Bolster, P., Trahan, E., 2009. Investing in Mad Money: Price and style effects. Financial Services Review 18 (1), 69–86.

Busse, J., Green, T., 2002. Market efficiency in real time. Journal of Financial Economics 65, 415–437.

Chan, K., Chen, N., 1988. An unconditional asset pricing test and the role of firm size as an instrumental variable for risk. Journal of Finance 43, 309–325.

Chen, C., Diltz, J., Huang, Y., Lung, P., 2011. Stock and option market divergence in the presence of noisy information. Journal of Banking & Finance 35, 2001–2020.

Cooper, M., Roberto C. Gutierrez, J., Marcum, B., 2005. On the predictability of stock returns in real time. The Journal of Business 78, 469–499.

De Long, J., Shleifer, A., Summers, L., Waldmann, R., 1990. Positive feedback investment strategies and destabilizing rational speculation. Journal of Finance 45, 379–395.

Engelberg, J., Sasseville, C., Williams, J., 2006. Is the market mad? Evidence from Mad Money. Working Paper, Northwestern University, available at SSRN: http://ssrn.com/abstract=870498 (downloaded on 09/12/2006).

Engelberg, J., Sasseville, C., Williams, J., 2012. Market madness? The case of Mad Money. Management Science 58, 351–364.

Fama, E., 1970. Efficient capital markets: A review of theory and empirical work. The Journal of Finance 25, 383–417.

Fama, E., French, K., 1992. The cross-section of expected stock returns. The Journal of Finance 47, 427–465.

Frazzini, A., Israel, R., Moskowitz, T., 2012. Trading costs of asset pricing anomalies. Working Paper, New York University, available at SSRN: http://ssrn.com/abstract=2294498 (downloaded on 01/12/2015).

Hobbs, J., Keasler, T., McNeil, C., 2012. Short selling behavior and Mad Money. Financial Review 47, 65–89.

Jeegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of Finance 48, 65–91.

Karniouchina, E., Moore, W., Cooney, K., 2009. Impact of Mad Money stock recommendations: Merging financial and marketing perspectives. Journal of Marketing 73, 244–266.

Keasler, T., McNeil, C., 2010. Mad Money stock recommendations: market reaction and performance. Journal of Economics and Finance 34, 1–22.

Kerl, A., Walter, A., 2007. Market responses to buy recommendations issued by personal finance magazines: Effects of information, price-pressure, and company characteristics. Review of Finance 11, 117–141.

Kudryavtsev, A., 2013. Stock price reversals following end-of-the-day price moves. Economics Letters 118, 203–205.

Liang, B., 1999. Price pressure: Evidence from the “Dartboard” column. Journal of Business 72, 119–134.

Lim, B., Rosario, J., 2010. The performance and impact of stock picks mentioned on “Mad Money”. Applied Financial Economics 20, 1113–1124.

Malkiel, B., 2003. The Efficient Market Hypothesis and its critics. The Journal of Economic Perspectives 17, 59–82.

Metcalf, G., Malkiel, B., 1994. The Wall Street Journal contests: The experts, the darts, and the Efficient Market Hypothesis. Applied Financial Economics 4, 371–374.

Neumann, J., Kenny, P., 2007. Does Mad Money make the market go mad? The Quarterly Review of Economics and Finance 47, 602–615.

Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Ross, S., 2006. A neoclassical look at behavioral finance: A tale of two anomalies. Unpublished Conference Presentation, available at: http://alumweb.mit.edu/upload/NZ/Two_Anomalies_Talk_New_Jersey_MIT_Club_-17728.pdf (downloaded on 01/20/2015).

Seasholes, M., Wu, G., 2007. Predictable behavior, profits, and attention. Journal of Empirical Finance 14, 590–610

Shleifer, A., Vishny, R., 1997. The limits of arbitrage. The Journal of Finance 52, 35–55.

Timmermann, A., Granger, C., 2004. Efficient Market Hypothesis and forecasting. International Journal of Forecasting 20, 15–27.

Wright, D., 1994. Can prices be trusted? A test of the ability of experts to outperform or influence the market. Journal of Accounting, Auditing and Finance 9, 307–23.

Zawadowski, A., Andor, G., Kértesz, J., 2006. Short-term market reaction after extreme price changes of liquid stocks. Quantitative Finance 6, 283–295.

IV. Essay 3 – High Frequency Trading Intensifies Intraday Extreme Events in Stock Returns

Author

Martin Niemann
University of Augsburg, Faculty of Business and Economics, Institute of Statistics
Universitätsstrasse 16. 86159 Augsburg, Germany
martin.niemann@gmail.com

Status

Currently in resubmit in: Journal of Business Economics (formerly ZfB Zeitschrift für Betriebswirtschaft)

Presented on Jan 29, 2015
at the Institute of Statistics, University of Augsburg, to Prof. Dr. Yarema Okhrin and Prof. Dr. Dr. h.c. Bamberg.

Abstract

This study examines the impact of high frequency trading (HFT) on intraday extreme events in U.S. stock returns. Although previous studies have generally found that, in aggregate, HFT improves broad market quality measures, it has remained an open question of whether this is also the case during turbulent markets, or whether HFT actually contributes to turbulence. Indeed, researchers and market operators suggest numerous ways in which HFT might exacerbate extreme events, and empirical investigations have found situational evidence that HFT intensified selling pressure during the “Flash Crash” of May 6, 2010. Using intraday price and HFT activity data, we examine the impact of HFT on extreme intraday price moves in general. To identify causal effects, we use the introduction of Regulation NMS (summer 2007) and the SEC Naked Access Ban (winter 2011/2012) as instrumental variables in two natural experiments. Our results strongly suggest that HFT activity exacerbates intraday extreme events.

"In the present environment, where high frequency and algorithmic trading predominate […], liquidity problems are an inherent difficulty that must be addressed. Indeed, even in the absence of extraordinary market events, limit order books can quickly empty and prices can crash simply due to the speed and numbers of orders flowing into the market and due to the ability to instantly cancel orders."

- Joint CFTC-SEC Advisory Committee on Emerging Regulatory Issues, 2011. Recommendations Regarding Regulatory Responses to the Market Events of May 6, 2010.

1. Introduction

Over the past decade, high frequency trading (hereafter *HFT*[1]), a high-speed, low-latency subset of algorithmic trading, has come to dominate many securities markets, particularly equities. Recent estimates by TABB Group and Morgan Stanley suggest that between 50% and 80% of U.S. stock transactions now involve HFT[2]. Meanwhile, concern has grown among researchers, policy makers and market operators over the impact of HFT on market behavior.

Dramatic advances in information technology have made HFT feasible, while structural and regulatory changes have motivated its growth. Along the entire value chain of securities trading, computers and software have largely replaced formerly phone- and human-based processes. Automation has shifted the balance between humans and algorithms in the competition for the shortest-term profit opportunities from market-making and arbitrage. Before the advent of HFT, when the U.S. trading landscape consisted only of the NYSE, AMEX, and NASDAQ stock exchanges and the futures markets, this role fell to the smartest and fastest human traders, most of whom were specialists and dedicated market-makers. In the new millennium,

---

[1] Throughout this paper, *HFT* refers to the *activity* of high frequency trading

[2] See http://www.nytimes.com/2012/10/15/business/with-profits-dropping-high-speed-trading-cools-down.html and http://www.ft.com/intl/cms/s/0/da5d033c-8e1c-11e1-bf8f-00144feab49a.html, last accessed 01/29/2014.

regulatory changes such as Regulation NMS ("National Market System", see SEC, 2005) have allowed the number of alternative trading venues to multiply, and encouraged private investment in technology to interlink and integrate the national market system. In the resulting fragmented automated trading environment, high frequency trading algorithms have largely taken over the role of market-makers and arbitrageurs ensuring short-run market efficiency.

The competitive environment today is best characterized as an arms race of increasing algorithmic trading speed (Gai et al., 2013) and intelligence (Credit Suisse, 2012). The race for the lowest *latency* – the time to obtain and process new information and trade on it – is in the order of microseconds for HFT infrastructure already co-located at the exchanges to minimize delay from the speed of light[3] (Johnson et al., 2013). Facing a physical boundary to speed, the remaining source of competitive advantage is developing ever-smarter algorithms. Given the costs of establishing and maintaining an HFT operation, it is natural for these institutions to seek every legal advantage to generate profits. For instance, by acting as an automated market maker serving liquidity demand faster than the competition (Hagströmer & Nordén, 2013), being the fastest to analyze and trade on news (Scholtus et al., 2014), or by anticipating and trading ahead of large orders from slower-reacting investors (SEC, 2010).

How have these developments affected relevant stock market characteristics? Does HFT pass on cost-savings from its automated intermediation, or does it impose externalities on other market participants? Among both academics and practitioners,

[3] For instance, NASDAQ advertises its new "1G Ultra" co-location connectivity to be nine microseconds faster than existing offers (http://www.nasdaqtrader.com/Trader.aspx?id=colo, last accessed 01/29/2014).

there is no clear consensus. Findings vary widely depending on the observer and the characteristics measured. A small but fast-growing body of empirical papers investigates the impact of HFT on measures of market quality. Hendershott et al. (2011) and Boehmer et al. (2013) study U.S. and international stock markets and find that algorithmic trading and HFT activity improve liquidity, enhance price discovery via quotes, and improve market efficiency. The latter also find that HFT increases volatility relative to what could be expected from narrower bid-ask spreads (Jones, 2013), especially on high-volatility days. Hagströmer & Nordén (2013) drill down further by separating HFT into market-making and opportunistic strategies, showing that high frequency market-making mitigates intraday volatility. While all of the above find an aggregate positive effect of HFT on most measures of market quality in normal market conditions, they raise the question whether HFT is equally beneficial in turbulent markets (Hendershott et al., 2011). Since, unlike the specialists that they have largely replaced, high frequency traders are not actually required to provide liquidity, they may exacerbate problems during turbulent markets by withdrawing from trading, or even by switching from being net providers to net demanders of liquidity.

The "Flash Crash" of May 6, 2010 put this question in the spotlight, when a large institutional sell order induced high frequency traders to exit the market (Easley et al., 2011). Cascading selling led to a dramatic market-wide crash of over 5%, with some stocks trading down to mere pennies, followed by a rebound (Kirilenko et al., 2014). Again, the interpretations vary. Jones (2013) argues extreme events exemplify a generic feature of markets with intermediaries temporarily overwhelmed by liquidity demand. In contrast, other authors hold that HFT could exacerbate volatility through new mechanisms such as *predatory trading* (i.e. trading that induces and/or exploits the trading needs of other investors, see e.g., Jarrow & Protter, 2012) or *positive-feedback loops* from correlated HFT behaviors and strategies (e.g., Zigrand et al., 2011).

Does HFT systematically exacerbate intraday extreme events? This paper extends the literature with empirical evidence on the impact of HFT, focusing particularly on the most turbulent periods. Building on Zawadowski et al. (2006), the proposed measure of intraday extreme events extracts the largest up or down price changes from intraday data, and determines their extremeness by comparing the move size with stocks' typical volatility in the same time of the day. Effectively, it is a measure of tail risk in the form of "*X*-sigma events"[4]. The dataset combines intraday price and HFT activity data, and confirms the massive growth in HFT over the past decade. A key challenge to determining a causal effect of HFT activity is that it is a potentially endogenous choice by traders. For instance, their actions might depend on the exact market characteristic we study, in that it might be particularly attractive for HFT to trade temporarily volatile stocks, rather than causing this volatility. To address this challenge, we follow the methodology established by Hendershott et al. (2011) and use exogenous shocks to the level of HFT activity in an instrumental variable regression to identify causality. Previous studies have found Regulation NMS and the SEC Naked Access Ban (hereafter NAB) as significant drivers of growth and decline in HFT activity (Chung & Chuwonganant, 2012), which we confirm in this paper.

The results of our instrumental variable regression strongly suggest that HFT indeed causes more extreme intraday price moves. Results are robust to common

[4] For example, if a stock's expected 1-hour intraday standard deviation (*sigma*) at 11 am is 1%, and it registers a 10% price move in that period, this would be a *10-sigma event.*

control variables, specification changes, and across different periods. These findings contribute evidence to the debate on the benefits and costs of HFT. To investors and issuers, excess volatility and tail risk are costs. They increase the riskiness of holding a position. Transitory price moves impede risk management via Stop-Loss orders by increasing adverse selection (i.e., stocks are sold at temporarily depressed prices). Hedging costs and needs increase with volatility and tail risk. Liquidity and volatility are important factors in asset pricing; therefore, issuers may face lower share prices and difficult conditions for equity issuance (Boehmer et al., 2013) . In sum, more severe intraday extreme events imply that investors pay for HFT activity by incurring externalities from increased market fragility. This adds to our understanding of the impact of HFT on market quality.

The outline of this paper is as follows. Section 2 motivates in further detail the focus of this study on HFT's effect on intraday extreme events, by establishing an understanding of known HFT strategies and potential mechanisms by which HFT might dampen or amplify intraday price moves. Section 3 presents the data and variable definitions for this study. We describe the exogenous shocks and their effect on HFT in section 4. Section 5 discusses the instrumental variable regression results. Finally, section 6 concludes this study and derives implications for future research.

2. High Frequency Trading, Intraday Extreme Events, and Previous Work

This section explores the potential link between *what HFT does* and how it might *mitigate or exacerbate extreme events* in intraday stock returns, within the context of previous work in this area. We start by describing HFT behaviors on a *stand-alone* basis (section 2.1), before moving on to consider possible mechanisms of *trader* *interaction* – involving both HFT and non-HFT – in section 2.2. We thereby focus our attention on behaviors and mechanisms with a conceivable influence on extreme events.

2.1. HFT Behaviors and Strategies

While there is not yet a universally accepted definition of HFT, the SEC (2010) defines it as trading activity which is characterized by (1) using extraordinarily high-speed hard- and software to generate, route and execute orders, (2) using co-location and direct data feeds offered by exchanges to minimize latency, (3) establishing and liquidating positions in extremely short time-frames, (4) submitting numerous orders and cancelling them shorty after submission (i.e. high *order-to-trade ratio*), and (5) ending the trading day with a position as close to zero as possible.

Liquidity provision through market-making-like strategies is one of the primary activities of high frequency traders. With their superior capability to instantly reflect new information in their quotes and much lower cost of intermediation, they have largely taken over the role of supplying liquidity from human traders (e.g., Hendershott et al., 2011). Liquidity is an important determinant of *non-fundamental volatility* (SEC, 2005). A liquid market allows investors to execute large orders close to the current price without causing large price moves away from fundamental value. A liquid market is *robust*, meaning that transitory demand or supply shocks have little impact on prices. In contrast, an illiquid market is *fragile*, meaning that large orders have more price impact and thus increase non-fundamental volatility. Hence, *ceteris paribus*, an increase in the supply of liquidity dampens market volatility and decreases the likelihood of extreme events. Hagströmer & Nordén (2013) differentiate between market-making and opportunistic strategies in their empirical study of HFT on the

NASDAQ-OMX Stockholm market, and provide evidence that the majority of HFT is market-making in nature and indeed reduces volatility.

Algorithmic execution and order anticipation is another major activity of high frequency trading. Several studies document features of HFT behavior which counteract liquidity improvements and increase the price impact of orders. Kim & Murphy (2013) show that effective bid-ask spreads have not actually improved as much as traditional market quality measures would suggest, once decreasing trade sizes – due to the increasing need to slice large orders into small tranches – are taken into account. This has made *algorithmic execution* of large orders the norm. Rather than risking significant market impact by placing one large order, investors often rely on proprietary or agency execution algorithms to buy or sell piece-by-piece over time. This innovation in trading creates opportunities in its own right. According to Quantitative Services Group (2010), this increased parceling leaves a "footprint" which sophisticated pattern-recognition algorithms can detect. This enables *order anticipation strategies* where HFT picks up an algorithmic execution early in the process, quickly builds a position in the same direction, and then immediately closes it, either by transacting with the original institutional order, or shortly thereafter. This is similar to the well-known strategy of front-running orders and has two effects: first, HFT competes with institutional orders for liquidity, effectively increasing the price impact of institutional trades. Second, as order anticipators close out their positions, a portion of this price impact is reversed. Indeed, based on a large sample of actual algorithmic executions, Quantitative Services Group (2010) find a significant reversal pattern within five minutes after the original order's completion. In essence, this is non-fundamental volatility. Tong (2013) confirms that HFT leads to an increase in institutional trading costs and highlights the opportunistic nature of HFT liquidity provision.

Latency and structural arbitrage. HFT also exploits structural weaknesses in the market structure. McInish & Upson (2012) document that high frequency traders use their *latency* advantage from their direct exchange price feed to pick off slow liquidity demanders who observe only the much slower national price feed. Then, when regular traders submit an order, the quotes they intended to trade on have already vanished, resulting in inferior executions. In addition, extreme HFT message volume can exploit weaknesses in *market structure*. Egginton et al. (2013) find systematic instances of "quote-stuffing", i.e. very high rates of orders that are immediately cancelled, which could slow down exchange systems, create additional latency arbitrage opportunities and confuse other traders. Both strategies increase non-fundamental volatility.

2.2. Interaction Mechanisms involving HFT

So far, we have identified several opportunistic HFT strategies, which profit-maximizing HFT operators appear to employ, and which tend to increase volatility. Beyond that, *trader interaction* yields several potential mechanisms.

Predatory trading. Theoretical models suggest HFT could create a mispricing and induce other players to trade in the same direction, causing crashes and price spikes (e.g., Boehmer et al., 2013; Brunnermeier & Pedersen, 2005; Jarrow & Protter, 2012). Thereby, HFT can profit by trading ahead of the price moves it creates. That some high frequency operators in fact follow such strategies is documented unequivocally in a report by the algorithmic execution team of Credit Suisse (2012).

They expose instances "momentum-ignition"[5], "layering"[6], and other predatory practices. Credit Suisse's emphasis on equipping its own agency execution algorithms with targeted defenses against these practices underlines their economic relevance. Most importantly, "momentum-ignition" and "layering" influence non-HFT players' trades and can therefore affect prices on much larger timescales.

Positive-feedback processes from correlated strategies. The greatest risks, arise when *positive-feedback loops* take hold among a larger share of market participants. The "Flash Crash" of May 6, 2010 serves as the prominent example for market-wide self-feeding volatility in several empirical studies showing that HFT exacerbated the crash by turning their liquidity provision into liquidity demand. Absent affirmative obligations to provide liquidity at all times, HFT liquidity supply is opportunistic and may exit the market when it is most needed. Jones (2013) argues that positive-feedback processes have occurred long before the advent of HFT and are rather a generic feature of equity markets than an HFT-specific risk. He points out several historical market-wide crashes as examples. In the "Black Monday" crash of October 19, 1987, for instance, many market participants used similar portfolio insurance algorithms, whose correlated hedging orders overwhelmed market-makers and exchange systems. In general, whenever sudden liquidity shocks from correlated orders overwhelm the supply-side, regardless of the time frame, self-feeding price moves can arise. Hence, risks from HFT are essentially not new. However, Jones' generalization ignores one important structural difference caused by the sub-second time scale of HFT: trading speed is hitting the limits of physics. In a few

5 "Momentum-ignition" attempts to trigger other market participants to trade quickly and induce a price-move, e.g. by moving prices beyond recent high or low prices.

6 By entering large orders at several price points, "Layering" creates a false impression of demand or supply in the order book to motivate others to trade.

microseconds, computers can perform only a few computations. This limits the scope and kinds of information that an HFT algorithm can process, and it restricts the complexity and diversity of profit-maximizing reactions to that information. Hence, Johnson et al. (2013) hold that high frequency traders are therefore bound to come up with similar strategies. At sub-second speeds, human traders cannot intervene[7], leaving the sub-second space completely to HFT. For the first time, in contrast to all previous market episodes, it is now possible for securities markets to enter short-run "all-machine phases" unencumbered by human traders. The authors identify this as one reason for "mini-flash crashes" observed in stocks in second and sub-second time frames. Along this line of reasoning, Sowers et al. (2012) propose a model in which correlated HFT strategies – and volatility – arise only from their herding on the smallest time scales compared to everyone else. The assertion that HFT entails correlated strategies is confirmed empirically by Breckenfelder (2013) and Chaboud et al. (2013).

In sum, there are plausible mechanisms and empirical evidence that HFT could cause volatility and extreme events. Above all, correlated HFT strategies and their interactions with slower traders could cause positive-feedback processes that are not limited to the sub-second HFT realm, but lead to excess price moves in longer period. In subsequent sections, we explore whether this is a general phenomenon.

[7] Minimal human reaction times for the simple task of registering and reacting to a visual signal average about 250 milliseconds, see http://www.humanbenchmark.com/tests/reactiontime/index.php, which excludes additional time for more demanding tasks such as making a trading decision.

3. Data, Variables, and Descriptive Analysis

The analyses in this paper use daily panel data with observations on 1,465 liquid U.S. stocks from January 3, 2006 to August 28, 2013. The daily panel aggregates intraday price and HFT activity data (see section 3.1) to our main regression variables for *HFT activity* (see section 3.2) and *intraday extreme events* (see section 3.3). Finally, section 3.4 provides an initial descriptive analysis of the data.

3.1. Raw data

The raw intraday data combine 1-minute price[8] and HFT activity[9] datasets. For every stock-minute, the data contain *open*, *high*, *low* and *close* prices and *volume*, as well as a measure for HFT activity (hereafter *HFT flags*). Our HFT flag variable is defined as the number of seconds of each minute (i.e., a number between 0 and 60) in which a stock has more than 100 quote changes per trade per second on any of the U.S. exchanges in the consolidated TAQ (trades-and-quotes) feed.[10] This is similar to the high *order-to-trade ratio* flag regularly used in HFT definitions by scholars (e.g., Egginton et al., 2013) and regulators (see section 2.1). While we cannot directly observe HFT activity, it is highly unlikely for a quote-to-trade ratio above 100 to occur with only human traders entering orders.[11] The measure is even somewhat conservative, since it ignores instances of moderate algorithmic quoting activity.

[8] Intraday price data is from financial market data service Lenz & Partner (VWD Group), providing 1-minute prices for the period January 3, 2005 to August 28, 2013.

[9] HFT activity data is provided by NANEX, a data feed service provider and research firm, covering the period January 3, 2006 to August 28, 2013

[10] Quote changes comprise any order changes at the top of the order book, i.e. any changes in the best bid-ask prices or sizes. The exclusion of order change messages deeper in the book is unlikely to affect our results compared to studies such as Hendershott et al. (2011) and Scholtus et al. (2014) using the full message volume. Boehmer et al. (2013) has found the time series of both variants highly similar.

[11] Humans, with effective reaction times rarely below one second (Johnson et al., 2013), can hardly achieve this.

From the intraday dataset containing all stocks listed on U.S. exchanges at the end of the sample period, we use a sample of the most liquid U.S. stocks, whose annual average price and daily dollar-volume exceed $10 and $20 million, respectively, for at least one year during the sample period[12]. We exclude half-days around U.S. exchange holidays and focus on regular trading hours from 9:30 am to 4:00 pm. Requiring at least two years of data, removing ETFs, closed-end funds, and correlated dual share-classes, and excluding six stocks with longer trading halts yields the final sample of 1,465 relatively liquid U.S. stocks[13]. We complement this with firm-level variables[14], including the free float and total number of shares outstanding.

3.2. Proxy Variables for HFT Activity

Aggregating the intraday price and HFT data yields two HFT proxies: the raw daily count of HFT flags for stock *i* in period *t*, which we denote HFT_flags_{it}, and a trading-volume-adjusted version, which we denote $HFT\_volume\_rate_{it}$, defined as:

$$HFT\_volume\_rate_{it} = \frac{HFT\_flags_{it} * 1{,}000{,}000}{Share\_volume_{it}} \tag{1}$$

The reason for this adjustment is that the simple activity count may overstate HFT activity[15], since higher overall trading volume makes random occurrences of 100 quotes-per-trade per stock-second slightly more likely.[16] We can interpret this

[12] This introduces survivorship bias, which could overstate time series increases in stocks' liquidity, price, and size measures, while potentially biasing volatility downwards. However, firm and time fixed effects in a panel regression framework pick up these slow moving compositional changes.

[13] From the sample of 5128 stocks, filtering out ETFs, closed-end funds, trusts, acquisition companies and other special-purpose vehicles removes 462 securities, the minimum price/liquidity criteria remove 3099 stocks, and the sample length and quality checks (including dual share-classes) remove another 102.

[14] Provided by FINVIZ

[15] We follow Hendershott et al. (2011) in adjusting raw activity data by share volume to make sure we are not merely picking up contemporaneous changes in market liquidity.

[16] We also performed the analysis adjusting HFT_flags_{it} by dollar volume; however, it is highly correlated to $HFT\_volume\_rate_{it}$ and does not yield any additional insights.

volume-adjusted rate as a conservative measure of "HFT market share" in total market activity.

3.3. Detecting and Measuring Intraday Extreme Events (IEE)

An intraday extreme event (hereafter *IEE*) measure requires two components: First, a measure of expected intraday volatility to assess which intraday price moves are "extreme", and second, a way of measuring the largest up or down price move that occurred in a given intraday period.

For the intraday volatility benchmark, we follow (Zawadowski et al., 2006) to take into account the intraday profile of volatility. It is highest at market open, when market participants price in accumulated news from the overnight non-trading period, lower during the middle of the day, and rises somewhat toward the market close along with trading volume (e.g., Kalev et al., 2004). Hence, we start by estimating the daytime volatility for stock *i* on day *t* in minute *m* (1 to *M* = 390 for each minute in regular trading hours from 9:30 am to 3:59 pm) for a day-start-bounded lookback window of *k* minutes, which we set to a standard 60 minutes throughout this study:

$$vola\_dsb_{itm} = \sqrt{\sum_{\mathrm{m}'=\max(1,\mathrm{m}-\mathrm{k})}^{m} \frac{r_{it\mathrm{m}'}^{2}}{1+m-\max(1,m-k)}} \tag{2}$$

where $r_{it\mathrm{m}'}$ is defined as the close-to-close log. return for all minutes except the first of each trading day, which we replace by its intra-minute return $\log(close_m/open_m)$ to exclude overnight returns. Thus $vola\_dsb_{itm}$ provides an estimate of one-minute realized volatility for a rolling *k*-minute lookback window, which is capped by the day start boundary (if $m < k$) to exclude overnight and previous day returns. To obtain the longer-term volatility benchmark, as in

(Zawadowski et al., 2006), we calculate a d = 60 day moving average of $vola\_dsb_{itm}$ for each minute m of the day:

$$daytime\_vola\_benchmark_{itm} = \sum_{t'=t-d}^{t-1} \frac{vola\_dsb_{it'm}}{d} * \sqrt{1 + m - \max(1, m - k)} \tag{3}$$

and apply the scaling rule[17] for volatility to estimate expected volatility for each k-minute rolling lookback window during the trading day.

To measure the largest directional price movements in the corresponding lookback window, we define $max\_move\_dn_{itm}$ for stock i on day t, during minute m of the trading day, as $\log(P_{max}/P_m)$ in the lookback window, from the highest *high* (P_{max}) to the *low* price (P_m) in current minute m. To illustrate, consider the calculation at 10:00 am. Due to the day start boundary, the lookback window for detecting the peak price shrinks to the 30 minutes from 9:30 to 9:59 am. Assuming the price peak within this period occurred at 9:41, $max\_move\_dn_{itm}$ captures the move from the high of minute "9:41" to the low in minute "10:00 am". $max\_move\_up_{itm}$ is defined vice-versa.

Combining these two components, we define our intraday extreme event measure as:

$$IEE\_dn_{itm} = \frac{|max\_move\_dn_{itm}|}{daytime\_vola\_benchmark_{itm}} \tag{4}$$

[17] It is important to note that intraday returns violate the scaling rule's *iid* distribution assumption, particularly due to minimum tick-sizes of 1 cent. However, our focus on liquid higher-price stocks mitigates this, and most importantly, the fixed effects in panel regressions pick up any persistent stock-specific bias. Therefore, for a given stock, as long as its price level and trade frequency do not change substantially, this volatility benchmark remains sufficiently consistent over time.

and define IEE_up_{itm} analogously for upside moves. In essence, we obtain a normalized measure of extremeness, which we can interpret as an "X-sigma move". Finally, to construct a daily time series variable for use in daily panel regressions, we aggregate the two 1-minute time series IEE_up_{itm} and IEE_dn_{itm} to a single daily time series IEE_{it} by aggregating their daily extremes[18]:

$$IEE_{it} = \max_{1 \le m \le M}(IEE\_dn_{itm}) + \max_{1 \le m \le M}(IEE\_up_{itm}) \tag{5}$$

IEE_{it} is our dependent variable in regressions to measure the effect of HFT on intraday extreme events. It captures the two points per day – one for upside, one for downside moves – when an asset exhibits the largest directional moves compared to its typical volatility in the same period. Partial *up/down* variants, using only one of the two intraday time series, could be used to analyze upside and downside extremes separately. Still, the combination has valuable characteristics: First, it reduces bias from market-wide intraday up and down trends. Second, if a stock gyrates wildly in both directions during the day, it leads to even higher values. This is valuable for our aim to measure unusual intraday volatility that is unlikely to be driven solely by fundamentals, but at least in part arising from trading dynamics itself (e.g., Farmer & Joshi, 2002).

3.4. Descriptive Analysis

Summary statistics for these variables can be seen in Table IV-1, which summarizes the panel with two groups of variables. Panel A reports statistics on the main regression variables: the HFT proxies *HFT_flags* and *HFT_volume_rate* and the intraday extreme event variable *IEE* from Equation (5). Panel B reports summary

[18] We exclude values of $IEE\_up/down_{itm}$ near the daily open and close (the first 10 and the last 15 minutes of the trading day) to remove observations with too short lookback due to the day start bound, and to prevent news-driven moves from dominating the sample.

statistics on the explanatory variables we use to capture changes in market conditions. Following Boehmer et al. (2013) and Hendershott et al. (2011) we use: *Volatility*, defined as $\log(intraday\ high/intraday\ low)$, *1/Price*, *Size* (total market capitalization), and *Turnover*, defined as the daily share volume in percent of a company's free float. All variables are 99.9% winsorized, i.e. values beyond the 0.05/99.95% quantiles are set to the quantile values.

Table IV-1: Full Sample Summary statistics

Notes: This table presents summary statistics of annual means per stock for the full combined price and HFT activity sample from January 3, 2006 to August 28, 2013.

Panel A: Main Regressors		2006	2007	2008	2009	2010	2011	2012	2013*
HFT_flags	Mean	1.4	37.2	72.4	100.5	161.0	357.3	198.9	158.6
(# flags / day)	5%	0.0	0.4	2.1	5.4	6.3	9.2	10.8	6.8
	Median	0.0	20.2	30.0	34.0	53.5	117.3	79.0	73.2
	95%	5.8	125.2	285.4	448.5	698.7	1,541.1	766.1	593.7
HFT_volume_rate	Mean	0.4	31.9	36.2	61.2	90.6	148.9	118.4	109.5
(# flags / 1 million	5%	0.0	1.0	3.4	9.2	13.6	19.6	19.2	12.6
shares traded)	Median	0.0	20.9	18.6	31.9	59.2	101.6	82.1	76.1
	95%	1.9	98.8	114.1	181.7	279.5	398.2	315.4	274.3
IEE	Mean	4.72	5.16	5.30	4.04	4.48	4.82	4.41	4.48
(x times intraday	5%	3.99	4.51	4.75	3.61	3.87	4.18	3.84	3.89
volatility benchmark)	Median	4.69	5.10	5.26	4.02	4.49	4.85	4.42	4.50
	95%	5.54	6.03	5.94	4.51	5.02	5.33	4.87	4.98
Panel B: Control Variables for Market Conditions									
Volatility (daily range)	Mean	2.55	2.86	5.49	4.53	2.88	3.11	2.58	2.32
(percent)	Median	2.36	2.69	5.28	4.13	2.68	2.90	2.29	2.02
Price	Mean	36.96	42.60	35.31	26.70	34.05	38.81	40.49	46.97
($)	Median	31.25	35.69	28.50	21.59	28.20	31.38	32.52	38.38
Size (market cap.)	Mean	15.47	17.85	14.88	11.07	13.22	13.89	14.27	16.12
($ billions)	Median	4.43	4.81	3.79	2.90	3.66	4.22	4.15	4.97
Share turnover	Mean	0.96	1.17	1.46	1.38	1.15	1.26	1.14	1.09
(percent of float)	Median	0.68	0.76	1.12	1.03	0.86	0.95	0.85	0.81
N (# stocks in sample)		1,223	1,293	1,334	1,378	1,423	1,465	1,465	1,465

* data until August 28

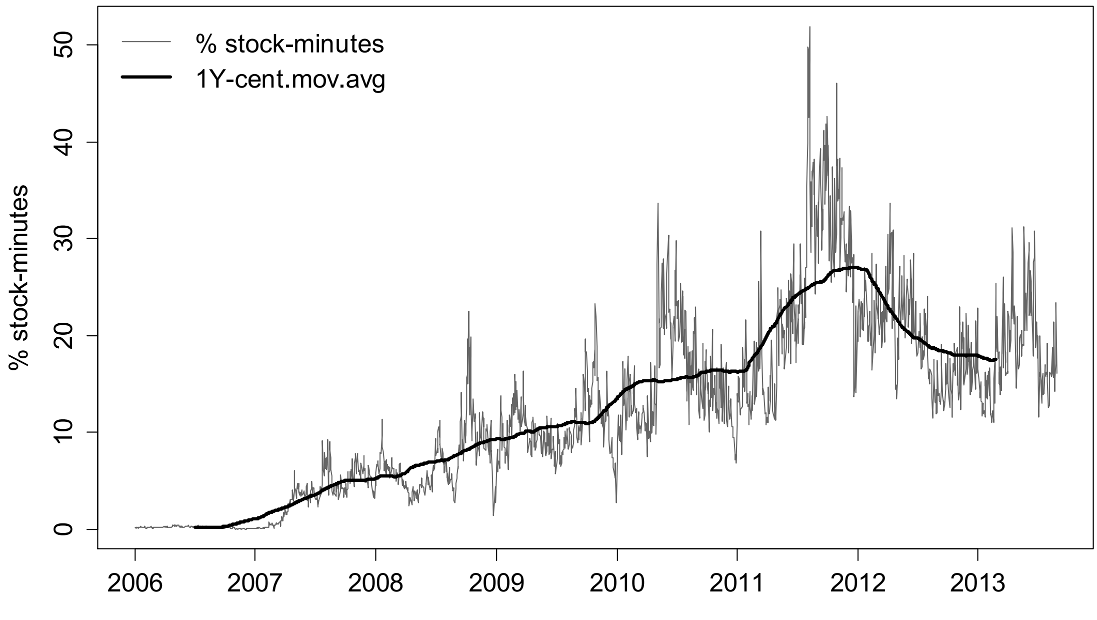
The HFT proxies in Panel A reveal a massive increase over time with substantial variation across stocks, both within and between market capitalization categories. Means are considerably larger than medians, suggesting a skewed distribution with concentrations of HFT activity in a small group of stocks. From 2008 onwards, once HFT has spread to almost all stocks, the 5% and 95% quantiles differ by a factor of roughly 100 for *HFT_flags*, and a factor of 20 for *HFT_volume_rate*, indicating the wide range of HFT activity levels across stocks. In contrast, there is much less cross-sectional variation in *IEE*, indicating the efficacy of normalizing by the stock's individual intraday volatility profile. Thus, the top and bottom quantiles differ only by a factor of 1.3 on average. Changes in market conditions, particularly in the financial crisis of 2008/2009, can also be seen in the control variables in Panel B: *Price* and *Size* drop considerably into 2009 before recovering substantially; *Turnover* and *Volatility* show an inverse pattern.

Figure IV-1: HFT Activity as Share of Total Stock-Minutes

Notes: This graph depicts the daily percentage of stock-minutes with at least one HFT quote activity flag and its 1-year (252 day) centered moving average. The derivation for each trading day t is

$$\%\,HFT\ minutes_t = \frac{\sum_{i=1}^{I_t} \sum_{m=1}^{M} \mathbf{1}_{HFT\_flags_{itm}>0}}{I_t M_t}$$

whereby i denotes individual stocks, I_t is the daily number of stocks in the dataset, m refers to the current minute in trading day (1 to M=390, i.e. each minute from 9:30 am to 3:59 pm EST).



The rising market penetration of HFT can be clearly observed in Figure 1, which effectively plots HFT market share, calculated as the daily percentage of HFT-affected stock-minutes. Growth starts in 2007 and persists until 2011, reaching daily peaks above 50% in 2011, when it affected more than half of the available "stock-time". Beginning in 2012, HFT activity drops notably and stabilizes at levels of about one-third below peak. In addition to these longer-term trends, HFT activity exhibits regular seasonal variations, including an annual drop of activity toward the end of the year.

Both the timing of entry and the eventual overall level of HFT activity varied considerably with market capitalization, as can be seen in Figure 2. HFT participants first targeted the large cap stocks, which is not surprising as these also tend to be the most liquid. As profit maximizing traders, they likely focused their initial algorithmic trading technology investments and trading capital on the stocks providing the most trading opportunities. Even after much of the entry period has passed in 2011 and 2012, large companies exhibit an order of magnitude more activity than small stocks.

Figure IV-2: HFT Flags per 1 Million Shares Traded, by Size Terciles

Notes: This chart shows daily averages and quantiles of HFT_volume_rate, defined as the daily number of HFT flags per 1 million shares trading volume. The y-axis is log-transformed. The daily mean is calculated across all stocks. All remaining time series represent one-year (252 day) centered moving averages of daily cross-sectional statistics.
Moving averages of daily means are given for Large/Medium/Small terciles by market capitalization, with firms ranked by their average market capitalization in the full period. The daily 5% and 95% percentiles indicate the upper and lower boundaries of distributions across stocks.

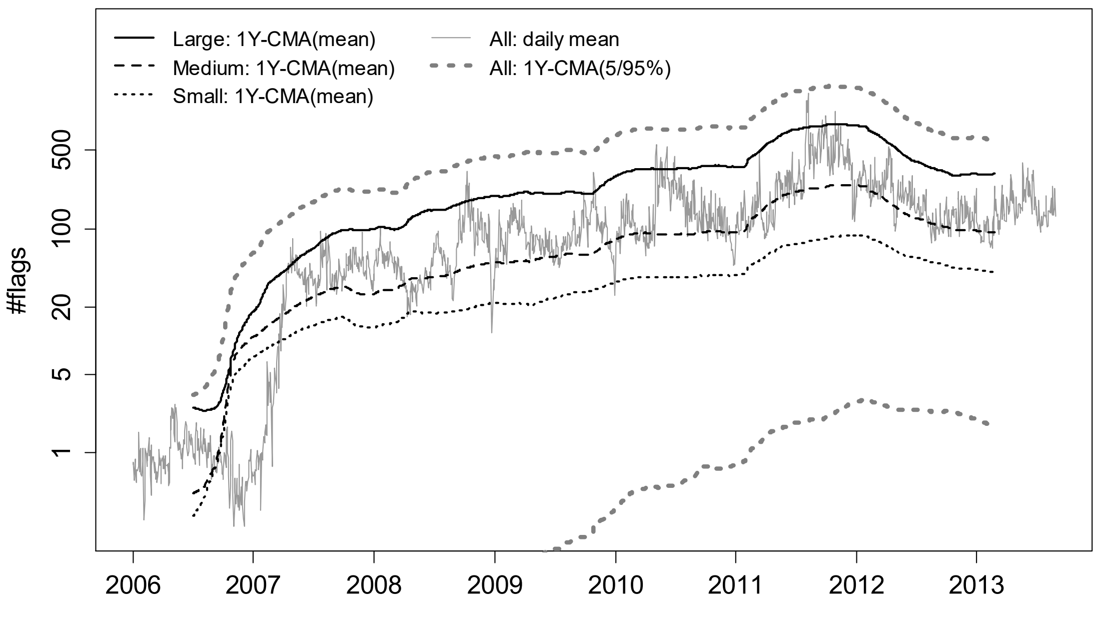
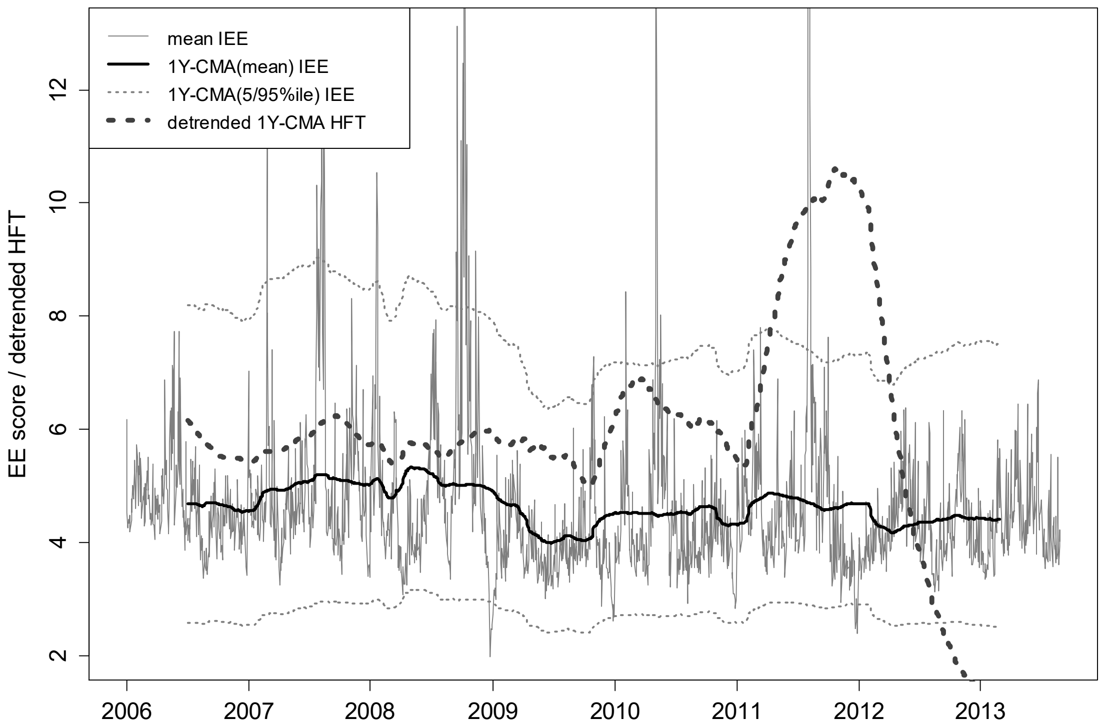


Figure 3 shows that increases in HFT activity coincide with increases in IEE. Peaks in extreme event severity also coincide with peaks in market volatility, (e.g. in the financial crisis of 2008/2009, and the "Flash Crash" of May 6, 2010), and one can also see seasonal patterns in IEE similar to those discussed above for HFT proxies. To make the comparison clearer, we compare IEE with a detrended version of the 1-year quasi-market-share time series from Figure 1. This adjusts for the strong growth in HFT activity over time, and allows us to focus on the residual variation around the trend. Viewed this way, HFT and IEE seem to be quite related. Their trends, peaks and troughs largely align, both for the mean and the outer quantiles (5/95%) of stocks' IEE values.

Figure IV-3: Intraday Extreme Events and Variations in HFT Quoting Activity

Notes: This chart displays aggregates of intraday extreme event (IEE) values and a detrended proxy of HFT quoting activity. It shows daily IEE means as well as one-year centered moving averages of the daily 5% and 95% percentiles of IEE.
The HFT proxy is defined as the one-year centered moving average of daily percentages of stock minutes with at least one HFT quote activity flag (see calculation in Figure 1). The detrended HFT proxy represents the residuals of a regression of the one-year average HFT proxy with a time-trend variable (rescaled to the value range of IEE in this figure).



From the initial analysis of HFT activity, we learn that it has grown substantially over time, both within stocks and in terms of market penetration. There are indications of a staggered rollout pattern in the market, with large, more liquid stocks being the first to exhibit significant levels of HFT, and temporary increases in HFT appear coincident with similar increases in IEE. It is not yet evident, however, whether this correlation reflects a causal relationship. Co-movement in IEE and HFT, by itself, does not allow inferring causality. Market participants might engage in HFT in a given stock because extreme events provide attractive trading opportunities. Both variables might be driven by variations in volatility and liquidity over time and across stocks. Nevertheless, this observation motivates the instrumental variable approach we take in the subsequent section.

4. Exogenous Shocks to HFT Activity

In econometric terms, the observed HFT proxy variables are endogenous to the system we are trying to analyze. This allows no causal inference for HFT’s effect on IEE in a standard regression. To disentangle the causal relationship between the variables, we follow the approach used in several previous studies on the effects of HFT on market characteristics: using an exogenous shock to HFT activity as an instrumental variable in a panel regression. The two shocks to HFT activity that we use as instruments were both actions by the SEC: the introduction of Regulation NMS and the Naked Access Ban (Chung & Chuwonganant, 2012).

4.1. The Rollout of Regulation NMS

On August 29, 2005, the SEC adopted Regulation NMS (hereafter *Reg. NMS*), a collection of rules which redefined the operations of U.S. equity markets with the goal of promoting liquidity, transparency and efficiency (SEC, 2005). As with many regulations, it did not have exactly the anticipated effect. Indeed industry observers such as TABB Group note:

"Reg. NMS started a progression of technology changes that […] prioritized speed over liquidity, […] fragmented the markets, drove ever-increasing messaging rates, created order-type complexity, and arguably enabled high frequency traders to take advantage of the very investors Reg. NMS was intended to protect, while actually making the markets less transparent for regulators." (Tabb, 2013)

Within Reg. NMS, the rule creating the most profound structural changes was *Rule 611*, the "Order Protection Rule", which established intermarket-protection against so-called trade-throughs, i.e., situations where one exchange would execute an order, despite a better price being available on another exchange. Implementing this protection effectively required *interlinking* all exchanges with one another to establish the "National Best Bid and Offer Price" (NBBO, see also SEC, 2005). Long geographical distances between exchanges and lags created by the process of aggregating all quotes to determine the current NBBO created the opportunity for HFTs. The SEC is well aware of this:

"Some proprietary firm's strategies may exploit structural vulnerabilities in the market or in certain market participants. For example, by obtaining the fastest delivery of market data through co-location arrangements and individual trading center data feeds, proprietary firms theoretically could profit by identifying market participants who are offering execution at stale prices." (SEC, 2010)

The ability to obtain prices faster than the rest of the public provided a significant economic incentive to invest and engage in HFT. McInish & Upson (2012) confirm this, showing the substantial profitability of NBBO-based latency arbitrage with an estimated profit of over $200 million annually for this strategy alone.

The implementation of Reg. NMS required substantial investments in market infrastructure, so that its final effective date was extended well into 2007. Full functional compliance with the main rules, including the "Order Protection Rule" was implemented in stages. 250 stocks (100 each from NYSE and NASDAQ, 50 from AMEX) were selected for a "Pilot Stocks Phase" from July 9 to August 20, 2007, with the rest of the market following thereafter ("All Stocks Phase"). For our purpose of investigating the effect of HFT on intraday extreme events, this staggered rollout is highly valuable. It constitutes a natural experiment, which allows comparing treated and untreated entities. In this case, the 96 "Pilot Stocks" in our sample can be compared to the rest of the population. Using a panel regression approach controls for changes in observed market conditions, as well as unobserved effects in individual days and firms, so that we can focus on the incremental effect from the exogenous shock to HFT from Reg. NMS.

If the economic incentives from Reg. NMS were indeed a positive shock to HFT, Reg. NMS might be a suitable instrument for our endogenous HFT proxy variables. A *valid instrument* has to satisfy the condition that it influences the dependent variable IEE only through its effect on HFT. This means that we have to be in a position to reasonably make the untestable assumption that (1) Reg. NMS does not influence IEE directly, (2) the reverse, i.e. IEE does not influence Reg. NMS, and that (3) there are no confounding Reg. NMS effects on IEE through other omitted variables. For requirement (2), it is helpful that the rollout schedule had been defined well in advance and also extended several times (SEC, 2007), so that we can confidently rule out an impact of IEE on the choice of stocks for the Reg. NMS rollout. For requirement (3), the stock and time fixed effects in a panel regression can compensate for biases, for instance, if particularly large or liquid stocks had been chosen for the "Pilot Stocks Phase". Finally, for requirement (1), we argue that, while there are well-defined technological changes and economic incentives for HFT traders to increase their activity, it is tough to make a case for other market participants changing their behavior significantly in a way that would influence volatility.[19] Hence, we have good reasons to assume that Reg. NMS constitutes a valid exogenous instrument for HFT.

4.2. The SEC Naked Access Ban

On November 30, 2011, the SEC "Naked Access Ban" (hereafter NAB) took effect. The rule banned a practice by broker/dealers providing their customers unfiltered ("naked") market access without any pre-order checks and monitoring of exposures (SEC, 2011). The practice of bypassing brokers' risk management checks allowed saving valuable fractions of seconds and was therefore heavily used by HFT[20]. Most importantly, the lack of monitoring and pre-order checks – with regard to exposures that could arise from the orders sent by the client – enabled HFT strategies involving high quote-volume order and cancellation sequences.

By banning this practice, the SEC substantially limited the scope to send an unlimited number of orders directly to exchanges. Representing an expected negative shock to HFT, the NAB provides an opportunity to study the effect of HFT trading on intraday extreme events. Better still, it covers a very different time period and market

[19] See Hendershott et al. (2011) for a thorough discussion of the requirements in context of a panel regression framework

[20] For a more detailed analyses of the rules and mechanisms involved, see Chakrabarty et al. (2014)

environment (2011, post financial crisis) compared to the time of the Reg. NMS rollout. Finally, analyzing a negative exogenous shock to HFT, undoing some of the growth spurred by Reg. NMS, provides a valuable validation opportunity.

Again, we have to be able to make the requisite assumptions for NAB to constitute a valid instrument for HFT. First, NAB should not influence IEE directly but only through HFT. This seems likely as the regulation directly targeted unfiltered market access heavily used by HFT, whereas for normal investors, naked access does not provide operational benefits. Second and inversely, that IEE has not influenced NAB seems reasonable given the SEC communication of its adoption well in advance. Finally, that other omitted variables with an influence on IEE confound our results at exactly the same time as the NAB introduction date seems unlikely, and, in case unobserved effects vary across stocks, the panel regression stock fixed effects can adjust for this. NAB has also been used as an exogenous instrumental variable for HFT activity in a recent working paper by Chakrabarty et al. (2014), and for our instrumental variable regressions involving NAB, we also borrow from their regression specification.

5. Empirical Findings

To investigate the effect of HFT on intraday extreme events, we use an instrumental variables approach, relying on the introduction of Reg. NMS and the Naked Access Ban as instruments. We construct subsamples for the periods in which the exogenous factors are active. In line with Hendershott et al. (2011), we exclude stocks with incomplete data to obtain a balanced panel, use standard price filters (stocks between $5 and $1000 during the sample period), and exclude small cap stocks below $250 million average market capitalization. This yields 1,197 stocks for the *Regulation NMS sample*, and 1,374 stocks for the *Naked Access Ban sample*, respectively[21]. The Regulation NMS "Pilot Stocks Phase" spanned the period from July 9 to August 20, 2007, and we follow Hendershott et al. (2011) and extend the sample by two months before and after, resulting in a period from May 9 to October 22, 2007. We construct the second sample for the Naked Access Ban in similar fashion, resulting in a period from October 3, 2011 to January 31, 2012, centered on the November 30, 2011 effective date for the ban. To keep track of the numerous different events, phases, and sample periods, Figure 4 provides a convenient overview.

Figure IV-4: Timeline of Events, Phases and Sample Periods

Notes: This graph summarizes important events and phases referenced throughout the paper, and the start and end dates of the data and subsamples. Abbreviations 2M and 6W refer to periods of 2 months and 6 weeks, respectively.

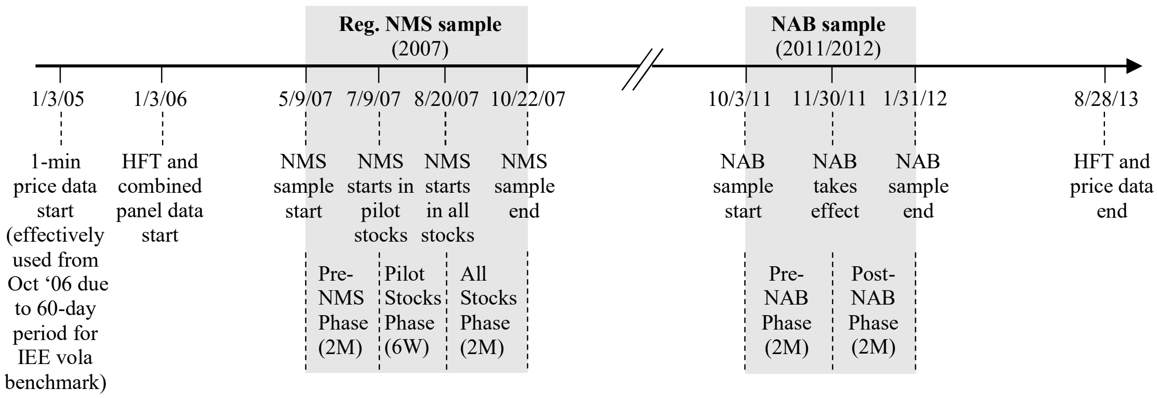


Table 2 summarizes the data for the two market phases covered by the subsamples. Compared to the Reg. NMS period in 2007, the level of HFT activity is much higher in the 2011/12 NAB sample, whereas the overall level of IEE is somewhat lower, possibly reflecting a calmer post-crisis market environment.

[21] For the *Reg. NMS sample* universe of 1,293 stocks with data in 2007, 72 stocks are removed due to missing data in the sample period, 19 are excluded due to price, and five due to size. For the *NAB sample* we start with 1,465 stocks with data in 2011, 10 stocks thereof are removed due to missing data, 76 due to price and five due to size.

Table IV-2: Subsample Summary Statistics

Notes: This table provides an overview of the two regression sample periods for the main regression variables. Following the approach for the regressions, we calculate statistics for the full sample and size terciles (Large, Medium, Small) by average market capitalization in the respective periods. "Within" standard deviation indicates the time series variation, averaged across stocks. "Across" standard deviation shows the variation of period means in the cross-section. The balanced panel samples for Regulation NMS (Panel A) and Naked Access Ban (Panel B) subperiods contain 1,197 and 1,374 stocks, respectively.

		IEE (*x-times id. vola.*)	HFT_flags (*# flags / day*)	HFT_vol.rate (*# flags / 1m shares*)	Volatility (*percent*)	Price (*$*)	Size (*$ billions*)	Turnover (*percent of float*)
Panel A: NMS Sample - *May 9 to October 22, 2007*								
All stocks	Mean	5.42	49.1	38.6	2.84	45.10	17.98	0.95
	Median	4.59	8.0	6.3	2.37	36.81	5.06	0.63
	Std Within	3.11	61.3	65.7	1.58	3.60	1.25	0.56
	Std Across	0.56	83.7	50.4	0.87	41.49	38.70	0.84
Large	Mean	5.32	97.4	49.9	2.44	59.99	46.60	0.64
	Median	4.58	35.0	13.2	2.07	48.89	24.61	0.44
Medium	Mean	5.43	34.7	38.0	2.79	44.08	5.56	0.92
	Median	4.60	6.0	6.0	2.32	37.75	5.06	0.63
Small	Mean	5.51	15.2	27.8	3.29	31.24	1.77	1.31
	Median	4.60	1.0	1.8	2.74	27.50	1.71	0.89
Panel B: NAB Sample - *October 3, 2011 to January 31, 2012*								
All stocks	Mean	3.92	497.7	222.0	3.18	39.69	14.20	1.19
	Median	3.59	107.0	91.5	2.70	31.28	4.15	0.83
	Std Within	1.55	339.4	202.5	1.48	2.76	0.82	0.65
	Std Across	0.33	807.7	363.7	1.16	42.25	30.60	1.03
Large	Mean	3.82	1,065.1	328.3	2.50	53.69	36.52	0.85
	Median	3.55	529.0	174.0	2.18	42.22	19.02	0.66
Medium	Mean	3.91	313.1	197.4	3.11	39.26	4.59	1.24
	Median	3.59	95.0	90.4	2.65	32.61	4.15	0.88
Small	Mean	4.01	115.0	140.3	3.92	26.13	1.49	1.48
	Median	3.62	25.0	47.8	3.38	21.65	1.46	1.01

5.1. Effect of Exogenous Shocks on HFT Activity

To test for effects of Reg. NMS on HFT and control variables as dependent variables Dep_{it} we use the following specification:

$$Dep_{it} = \alpha_i + \gamma_t + \beta NMS_{it} + \varepsilon_{it} \qquad (6)$$

where α_i and γ_t are stock and day fixed effects, and NMS_{it} is the Regulation NMS introduction dummy variable. For each stock i, NMS_{it} is zero for all days before the stock-specific rollout date and one thereafter. Including day-specific fixed effects removes any time trends in dependent variables, and effectively compares the change in Dep_{it} for Reg. NMS pilot stocks with that of the non-pilot stocks.

We use a similar specification for the Naked Access Ban period, replacing the γ_t day fixed effect with $Trend_t$, a linear time trend variable, due to the absence of a staggered rollout in NAB_t, which would allow us to separately identify day-specific fixed effects:

$$Dep_{it} = \alpha_i + \gamma Trend_t + \beta NAB_t + \varepsilon_{it} \tag{7}$$

The results for these regressions strongly suggest that Reg. NMS and the Naked Access Ban substantially influenced HFT activity, with Reg. NMS leading to a substantial increase and the NAB leading to a substantial decrease. This can be seen in Table 3, where the coefficient on the NMS_{it} dummy for all stocks indicates that the introduction of Reg. NMS increases HFT flags by about 40 per day, an increase of 80 percent over the overall sample mean of roughly 50 (see Table 2).

Table IV-3: Reg. NMS Rollout Effect on HFT Proxy and Control Variables

Notes: This table shows the effect of the Regulation NMS introduction on HFT proxies and other covariates used in this study. The model specification uses each variable as dependent variable Dep_{it} in the reduced form 1st stage regression (Equation 6):

$$Dep_{it} = \alpha_i + \gamma_t + \beta NMS_{it} + \varepsilon_{it}$$

where α_i and γ_t are stock and day fixed effects, and NMS_{it} is the Regulation NMS introduction dummy variable (zero for all days before the stock-specific rollout date and one thereafter). We run the regression for all stocks as well as for size terciles (Large, Medium and Small) by stocks' average market capitalization in the sample period. The sample covers the period from May 9 to October 22, 2007, which includes the Regulation NMS rollout period from July 9 to August 20, 2007, plus/minus two trading months.

	All stocks	Size Terciles		
Reg. NMS Effect (β)		Large	Medium	Small
on HFT proxies:				
HFT_flags	42.6 ***	37.2 ***	38.6 ***	12.2 **
log. HFT_flags	0.29 ***	0.27 ***	0.20 ***	0.17 **
log. HFT_vol.rate	0.32 ***	0.24 ***	0.15 **	0.40 ***
on Control variables:				
Volatility	-0.090 *	0.123 *	0.033	-0.415 ***
1/Price	0.000 **	0.000 ***	0.001 **	-0.001 **
log. Size	-0.006 **	-0.006 **	-0.007	-0.001
Turnover	0.000 **	0.000 ***	0.001 **	-0.001 **

Significance levels: *** <.001, ** <.01, * <.05
observations: 1,197 * 114 (stocks * days)

The log-transformed HFT proxies, which reduce the impact of the substantial absolute differences in HFT activity levels across stocks, still imply a 30% increase for both HFT activity measures. All significance estimates are based on standard errors double clustered across firms and time[22]. By contrast, coefficients on control variables are inconsistent and economically insignificant. Taken together, it is clear that Reg. NMS was a significant positive shock to HFT.

Table IV-4: SEC Naked Access Ban Effect on HFT Proxy and Control Variables

Notes: This table shows the effect of the SEC Naked Access Ban on HFT proxies and other covariates used in this study. The model specification uses each variable as dependent variable Dep_{it} in the reduced form 1st stage regression (Equation 7):

$$Dep_{it} = \alpha_i + \gamma Trend_t + \beta NAB_t + \varepsilon_{it}$$

where α_i is a stock fixed effect, $Trend_t$ is a linear time-trend variable, and NAB_t is the Naked Access Ban dummy variable (zero prior to November 30, 2011 and 1 thereafter). Due to the lack of cross-sectional variation in NAB_t, $Trend_t$ replaces the day fixed effect (see specification in Table 3) to account for potential slow-moving variation in other unobserved variables that might influence the dependent variables. We run the regression for all stocks as well as for size terciles (Large, Medium and Small) by stocks' average market capitalization in the sample period. The sample covers the period from October 3, 2011 to January 31, 2012, which includes the SEC Naked Access Ban (NAB) rollout day (November 30, 2011) plus/minus 2 months.

	All stocks	Size Terciles			All stocks
NAB Effect (β)		Large	Medium	Small	**Trend (γ)**
on HFT proxies:					
HFT_flags	-166.8 ***	-362.3 ***	-82.7 ***	-55.5 ***	-2.86 ***
log. HFT_flags	-0.24 ***	-0.29 ***	-0.18 ***	-0.25 ***	-0.01 ***
log. HFT_vol.rate	-0.23 ***	-0.27 ***	-0.15 ***	-0.28 ***	-0.01 ***
on Control variables:					
Volatility	0.322 ***	0.267 ***	0.278 ***	0.420 ***	-0.030 ***
1/Price	0.000 ***	0.000 **	-0.001 **	0.000 *	0.000 ***
log. Size	-0.055 ***	-0.046 ***	-0.056 ***	-0.064 ***	0.002 ***
Turnover	0.000 ***	0.000 **	-0.001 **	0.000 *	0.000 ***

Significance levels: *** <.001, ** <.01, * <.05
observations: 1,374 * 82 (stocks * days)

By contrast, the Naked Access Ban substantially decreased HFT, as seen in Table 4, reducing HFT activity by roughly one third, similar to the findings of Chakrabarty et al. (2014). The control variables exhibit economically larger and more significant coefficients on NAB_t than in the previous analysis on NMS_{it}. This is due to the fact that we cannot use day fixed effects in the regression equation (7), so that deviations above and below the linear time trend can show up as significant coefficients for NAB_t. For example, for the NAB effect on $Volatility$, the point

[22]To estimate regressions with a high number of factor levels, we use the fixed effects implementation in R by Gaure (2013).

estimate of 0.32 for the full sample implies below-trend volatility before the effective date of the ban and above-trend volatility thereafter, but not necessarily higher absolute volatility[23].

5.2. Effect of HFT Activity on Intraday Extreme Events

To identify causal effects from HFT on intraday extreme events, we use the staggered introduction of Regulation NMS and the introduction of the Naked Access Ban as exogenous instruments for HFT proxies in the two subsamples. The dependent variable is the square root of *IEE* measures as defined in equation (5) and we use log-transformed HFT proxy variables[24] for the following regression specification in the *Regulation NMS sample*:

$$\sqrt{IEE_{it}} = \alpha_i + \gamma_t + \beta \log(1 + HFT_{it}) + \delta X_{it} + \varepsilon_{it} \tag{8}$$

where α_i and γ_t are stock and day fixed effects, HFT_{it} is the selected HFT proxy variable (HFT_flags_{it} or $HFT\_volume\_rate_{it}$) and X_{it} is a vector of control variables ($Volatility_{it}$, $1/Price_{it}$, $Size_{it}$, and $Turnover_{it}$). Note that in the instrumental variable approach HFT_{it} is first regressed against the Reg. NMS dummy HFT_{it} and all other explanatory variables to determine the exogenous source of variation in HFT. Price and size are lagged by one period, but in contrast to Boehmer et al. (2012), we use contemporaneous volatility and turnover, due to their high time series variance in order to ensure that they accurately control for the expected level of

[23]The $\gamma = -0.03$ coefficient on the linear trend variable *Trend*$_t$ suggests that on average, pre-NAB volatility is 1.2 percent higher than post-NAB volatility (40 trading days average difference * 0.03).

[24]The economic rationale comes from the intuition that IEE captures price shocks reflecting the market impact of large order imbalances. Models for transitory market impact (e.g., Almgren, 2005) follow concave functions of relative volume with exponents around 0.5 for the dependent variable. In addition, residual distributions are more symmetrical for $\sqrt{IEE_{it}}$ compared to $\log(IEE_{it})$. In any case, regression results are robust against changing the transformations of both IEE and HFT variables.

IEE intensity given potential news shocks and resulting trading activity on the same day.

Table IV-5: Effect of HFT Activity on Daily Intraday Extreme Events using Regulation NMS Rollout as Instrument for HFT

Notes: This table shows the regression results for daily intraday extreme events (IEE_{it}) on two proxies for HFT activity (HFT_{it}). The sample covers the period from May 9 to October 22, 2007, which includes the Regulation NMS rollout period from July 9 to August 20, 2007, plus/minus two trading months. We run the regressions for all stocks as well as for size terciles (Large, Medium and Small) by stocks' average market capitalization in the sample period. For all regressions, we use the staggered rollout of Regulation NMS as instrument for the endogenous HFT proxy variables HFT_{it} to isolate the causal effect of HFT on IEE. The specification is (Equation 8):

$$\sqrt{IEE_{it}} = \alpha_i + \gamma_t + \beta \log(1 + HFT_{it}) + \delta X_{it} + \varepsilon_{it}$$

where α_i and γ_t are stock and day fixed effects, HFT_{it} is the selected HFT proxy variable (HFT_flags_{it} or $HFT\_volume\_rate_{it}$) and X_{it} is a vector of control variables ($Volatility_{it}$, $1/Price_{it}$, $log(Size_{it})$ and $Turnover_{it}$). Panel A regresses IEE_{it} on HFT_flags_{it}, the daily number of seconds per stock with at least 100 quotes per trade, Panel B regresses IEE_{it} on $HFT\_volume\_rate_{it}$, the daily number of HFT activity flags per 1 million shares traded. The right-most column of each panel reports the coefficients in the 1st-stage of the 2-step instrumental variable regression, which regresses HFT_{it} on NMS_{it} and controls. In addition, F-statistics are shown for a Wald-Test of the significance for NMS_{it} entering the 1st stage regression.

Panel A:	*IEE on HFT_flags*							*HFT_flags on NMS*
		log. HFT_flags	Volatility	1/Price	log. Size	Turnover	(1st stage regression)	
All stocks	Coeff.	0.121	0.196	0.875	0.021	3.357	Coeff.	0.293
	Pr(>|t|)	0.000 ***	0.000 ***	0.172	0.437	0.000 ***	F-stat.	112.08
Large	Coeff.	0.175	0.226	4.514	-0.065	3.415	Coeff.	0.263
	Pr(>|t|)	0.000 ***	0.000 ***	0.000 ***	0.248	0.000 ***	F-stat.	58.06
Medium	Coeff.	0.021	0.205	-4.720	0.022	5.906	Coeff.	0.191
	Pr(>|t|)	0.843	0.000 ***	0.000 ***	0.735	0.001 ***	F-stat.	(9.64)
Small	Coeff.	-0.134	0.186	-1.229	0.172	7.451	Coeff.	0.214
	Pr(>|t|)	0.161	0.000 ***	0.416	0.001 ***	0.000 ***	F-stat.	12.49
Panel B:	*IEE on HFT_volume_rate*							*HFT_v.rate on NMS*
		log. HFT_vol.rate	Volatility	1/Price	log. Size	Turnover	(1st stage regression)	
All stocks	Coeff.	0.108	0.200	0.311	0.002	6.126	Coeff.	0.328
	Pr(>|t|)	0.000 ***	0.000 ***	0.587	0.944	0.000 ***	F-stat.	116.76
Large	Coeff.	0.177	0.235	4.168	-0.079	10.405	Coeff.	0.260
	Pr(>|t|)	0.000 ***	0.000 ***	0.001 ***	0.185	0.000 ***	F-stat.	54.95
Medium	Coeff.	0.024	0.206	-4.702	0.017	6.567	Coeff.	0.168
	Pr(>|t|)	0.843	0.000 ***	0.000 ***	0.851	0.000 ***	F-stat.	(6.36)
Small	Coeff.	-0.070	0.182	-0.006	0.174	5.099	Coeff.	0.410
	Pr(>|t|)	0.132	0.000 ***	0.994	0.000 ***	0.000 ***	F-stat.	32.58

Significance levels: *** <.001, ** <.01, * <.05 ( ): Wald F-stat < 10
observations: 1,197 * 114 (stocks * days)

Table 5 reports the results of the instrumental variable regression, using HFT_flags_{it} (Panel A) and $HFT\_volume\_rate_{it}$ (Panel B) as HFT proxies. We find an economically and statistically significant effect of HFT activity on intraday extreme events for large stocks but not for smaller stocks. The HFT regression coefficients of 0.18 for large stocks imply that a doubling of HFT activity compared to trading volume leads to an increase in IEE_{it} of 0.6, or the largest price moves per day grow from 5.3 to "5.9-times sigma", which represents a more than 10% increase in severity[25]. This example is just for a doubling in HFT activity. Yet the actual time series variation of HFT activity easily reaches factors of 5 to 10 for daily data and several orders of magnitude intraday. As a result, bursts of HFT activity could have much larger effects.

The right column in Table 5 also shows the coefficients and the F-Test for instrument NMS_{it} entering the 1st stage regression for each model. The F-values show clearly that the instrument is much weaker for medium and small stocks. This lack of statistical power mirrors the results by Hendershott et al. (2011) who argue that their instrument might be weaker for smaller stocks because of less time series variation in these groups. Indeed, the 96 stocks in our sample belonging to the NMS pilot stocks per July 9, 2007, are concentrated in large stocks (49 stocks), with medium and small terciles relatively underrepresented (19 and 18 pilot stocks, respectively). Beyond that, we know that HFT activity is an order of magnitude higher for large stocks than for smaller stocks, so that the effective HFT-increase in small stocks due to Reg. NMS is economically tiny in comparison. Finally, it also fits to our insight from descriptive analysis in section 3.4: when resource-constrained traders invest in additional HFT capacity, they are more likely to apply it first to large stocks offering the greatest profit opportunities. In sum, the concentration of HFT-effects in large stocks fits well to economic and statistical circumstances.

[25] Calculation: using the IEE sample average of 5.3 for large companies (see table 2), we get $\left(\sqrt{5.3} + 0.18 * \log(2)\right)^2 = (2.3 + 0.12)^2 = 5.3 + 0.6$, where 0.6 is the HFT effect.

The findings from the NAB sample confirm these results. The specification for the Naked Access Ban regressions differs from the Reg. NMS regressions only by replacing time fixed effects γ_t by $Trend_t$.and NMS_{it}.by NAB_t.as instrument for HFT. The results indicate that HFT activity has a substantial impact on intraday extreme events, not just for large stocks but all size groups. Full sample coefficients on HFT are roughly three times larger than in the Reg. NMS period. The economic rationale for the increase in significance for medium and small market capitalization stocks has two potential drivers. First, the intensity of HFT activity in these stocks is an order of magnitude higher in 2011 than it was in 2007. Second, the NAB effect does not depend on resource prioritization decisions by HFT firms, since the ban limits order submission in general.Table 6 reports the regression results using NAB_t as an instrument for HFT proxy variables. We follow Chakrabarty et al. (2014) and use the following specification, derived from equation (8):

$$\sqrt{IEE_{it}} = \alpha_i + \gamma Trend_t + \beta \log(1 + HFT_{it}) + \delta X_{it} + \varepsilon_{it} \qquad (9)$$

The specification for the Naked Access Ban regressions differs from the Reg. NMS regressions only by replacing time fixed effects γ_t by $Trend_t$ and NMS_{it} by NAB_t as instrument for HFT. The results indicate that HFT activity has a substantial impact on intraday extreme events, not just for large stocks but all size groups. Full sample coefficients on HFT are roughly three times larger than in the Reg. NMS period. The economic rationale for the increase in significance for medium and small market capitalization stocks has two potential drivers. First, the intensity of HFT activity in these stocks is an order of magnitude higher in 2011 than it was in 2007. Second, the NAB effect does not depend on resource prioritization decisions by HFT firms, since the ban limits order submission in general.

Table IV-6: Effect of HFT Activity on Daily Intraday Extreme Events using SEC Naked Access Ban as instrument for HFT

Notes: This table shows the regression results for daily intraday extreme events (IEE_{it}) on two proxies for HFT activity (HFT_{it}). The sample covers the period from October 3, 2011 to January 31, 2012, which includes the SEC Naked Access Ban (NAB) rollout day (November 30, 2011) plus/minus 2 months. We run the regressions for all stocks as well as for size terciles (Large, Medium and Small) by stocks' average market capitalization in the sample period. For all regressions, we use the Naked Access Ban as instrument for the endogenous HFT proxy variables HFT_{it} to infer the causal effect of HFT on IEE. The specification is (Equation 9):

$$\sqrt{IEE_{it}} = \alpha_i + \gamma Trend_t + \beta \log(1 + HFT_{it}) + \delta X_{it} + \varepsilon_{it}$$

where α_i is a stock fixed effect, $Trend_t$ is a linear time-trend variable, HFT_{it} is the selected HFT proxy variable (HFT_flags_{it} or $HFT\_volume\_rate_{it}$) and X_{it} is a vector of control variables ($Volatility_{it}$, $1/Price_{it}$, $log(Size_{it})$ and $Turnover_{it}$). Panel A regresses IEE_{it} on HFT_flags_{it}, the daily number of seconds per stock with at least 100 quotes per trade, Panel B regresses IEE_{it} on $HFT\_volume\_rate_{it}$, the daily number of HFT activity flags per 1 million shares traded. The right-most column of each panel reports the coefficients in the 1st-stage of the 2-step instrumental variable regression, which regresses HFT_{it} on NAB_t and controls. In addition, F-statistics are shown for a Wald-Test of the significance for NAB_t entering the 1st stage regression.

Panel A:	*IEE on HFT flags*							*HFT_flags on NAB*	
		log. HFT_flags	Volatility	1/Price	log. Size	Turnover	Trend	(1st stage regression)	
All stocks	Coeff.	0.487	0.117	5.545	-0.170	4.724	0.006	Coeff.	-0.208
	Pr(>|t|)	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	F-stat.	96.68
Large	Coeff.	0.437	0.154	1.411	-0.400	7.998	0.007	Coeff.	-0.264
	Pr(>|t|)	0.000 ***	0.000 ***	0.226	0.000 ***	0.000 ***	0.000 ***	F-stat.	31.40
Medium	Coeff.	0.658	0.115	9.695	0.018	5.448	0.010	Coeff.	-0.156
	Pr(>|t|)	0.000 ***	0.000 ***	0.000 ***	0.772	0.000 ***	0.000 ***	F-stat.	34.98
Small	Coeff.	0.431	0.100	4.999	-0.150	4.278	0.004	Coeff.	-0.207
	Pr(>|t|)	0.000 ***	0.000 ***	0.000 ***	0.006 **	0.000 ***	0.000 ***	F-stat.	31.51
Panel B:	*IEE on HFT volume rate*							*HFT_v.rate on NAB*	
		log. HFT_vol.rate	Volatility	1/Price	log. Size	Turnover	Trend	(1st stage regression)	
All stocks	Coeff.	0.531	0.137	4.708	-0.337	19.931	0.006	Coeff.	-0.191
	Pr(>|t|)	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	F-stat.	96.68
Large	Coeff.	0.498	0.184	-0.465	-0.581	29.033	0.007	Coeff.	-0.231
	Pr(>|t|)	0.000 ***	0.000 ***	0.746	0.000 ***	0.000 ***	0.000 ***	F-stat.	31.40
Medium	Coeff.	0.798	0.142	6.798	-0.302	30.447	0.011	Coeff.	-0.129
	Pr(>|t|)	0.000 ***	0.000 ***	0.000 ***	0.001 ***	0.000 ***	0.000 ***	F-stat.	34.98
Small	Coeff.	0.408	0.114	4.622	-0.216	14.497	0.004	Coeff.	-0.219
	Pr(>|t|)	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	F-stat.	31.51

Significance levels: *** <.001, ** <.01, * <.05 ( ): Wald F-stat < 10

observations: 1,374 * 82 (stocks * days)

In sum, the instrumental variable regressions strongly suggest that HFT activity exacerbates extreme intraday price moves, particularly for large stocks. The findings are consistent across very different market periods – summer 2007 and the end of 2011 – as well as different model specifications[26].

[26] In addition to the NAB validation, we have also tested models including the lagged dependent variable, different variable transformations (e.g. using raw *IEE*), and differently defined or lagged control variables. To check if intraday effects drive our results, we have also tested extended specifications using (1) intraday HFT activity measured *contemporaneously* to the respective daily extreme events, (2) contemporaneous turnover as control

6. Conclusions

This study investigates the impact of HFT activity on intraday extreme events in U.S. stock returns. This question is raised by the empirical literature, where several authors find largely positive HFT effects on aggregate market quality, but at the same time call into question HFT's role in turbulent markets. Moreover, market observers and researchers report aggressive HFT strategies, which might cause non-fundamental volatility. Beyond stand-alone HFT actions, researchers suggest interaction mechanisms by which HFT exacerbates extreme events. With several investigations of the "Flash Crash" of May 6, 2010, the literature provides situational evidence that high frequency traders – while not causing the crash – contributed to extreme downside volatility.

This paper adds empirical evidence to the HFT debate, focusing on the role of HFT in causing market turbulence. We measure intraday extreme events as the size of a stock's daily largest intraday up and down price moves compared to its typical intraday volatility during the same times of day. Using an instrumental variable approach with Regulation NMS (summer 2007) and the SEC Naked Access Ban (winter 2011/2011) as exogenous shocks to HFT, we show that increased activity by high frequency traders causes more severe intraday extreme events. Put differently, tail events become larger. This finding is not limited to single market events, but is robust across two widely different market periods – before and after the financial crisis of 2008/2009, before and after the massive growth in HFT market penetration – and for a positive and a negative exogenous shock to HFT, respectively. The results add to our understanding of the market quality effects of HFT, and relativize some of

variable and (3) time of day effects. For the sake of brevity, these regressions are not reported here since they deliver comparable results.

the positive effects of HFT found in previous studies. Overall, while there are positive effects of HFT, they are paid for at least in part by ordinary investors who have to accept more fragile market conditions.

Numerous mechanisms could be responsible for this empirical finding. HFT liquidity supply can vanish in fractions of a second if market conditions suggest it is profit maximizing for high frequency traders to exit the market. Furthermore, in pursuit of licit profit motives, some of the algorithms used by high frequency traders directly reduce market quality and increase trading costs for ordinary investors in numerous ways. As a category of market participants, they operate close to physical speed limits and therefore *must* follow correlated strategies. This introduces new positive-feedback mechanisms which could be responsible for larger tail events. To determine which subset of HFT behaviors or interaction mechanisms is responsible for the finding of more intense extreme events, is beyond the scope and data of this study but an important avenue for research. Tail risks and more generally non-fundamental volatility are costs to both investors and issuers, and hence, overall welfare implications of HFT are yet unclear.

For future research, the design of more robust market systems, and the conception of rules directing technological innovations in financial markets towards positive effects on market quality, is an open question. Assessing the welfare effects of HFT, as well as market design choices in response, requires more detailed analysis using full order book data and identifying the actions and impacts of individual HFT firms and strategies.

Acknowledgements

I would like to thank Prof. Günter Bamberg and Robert McMillan for helpful comments and discussions, Eric Hunsader and Nate Rock from NANEX for data and helpful discussions on high frequency trading activity, and Lenz & Partner (VWD Group) for providing intraday stock price data. Additional thanks go to Emily Niemann, Alexander Wiegelmann and an anonymous referee for helpful comments and suggestions. The usual disclaimer applies.

References (Chapter IV)

Almgren, R., 2005. Equity market impact. RISK (July 2005) 57–62.

Boehmer, E., Fong, K., Wu, J., 2013. International Evidence on Algorithmic Trading. Working Paper, EDHEC Business School.

Breckenfelder, H., 2013. Competition between High-Frequency Market Makers, and Market Quality. Working Paper, Stockholm School of Economics.

Brunnermeier, M., Pedersen, L., 2005. Predatory trading. Journal of Finance 60, 1825–1864.

Chaboud, A., Hjalmarsson, E., Vega, C., Chiquoine, B., 2014. Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. The Journal of Finance 69, 2045–2084.

Chakrabarty, B., Jain, P., Shkilko, A., Sokolov, K., 2014. Quote Intensity and Market Quality: Effects of the SEC Naked Access Ban. Working Paper, Saint Louis University.

Chung, K., Chuwonganant, C., 2012. Regulation NMS and Market Quality. Financial Management 41, 285–317.

Credit Suisse Trading Strategies, 2012. High Frequency Trading – Measurement, Detection and Response. Available at https://edge.credit-suisse.com/edge/Public/Bulletin/Servefile.aspx?FileID=23285&m=923226224, last accessed 02/07/2014.

Easley, D., López de Prado, M., O'Hara, M., 2011. The Microstructure of the "Flash Crash": Flow Toxicity, Liquidity Crashes, and the Probability of Informed Trading. The Journal of Portfolio Management 37, 118–128.

Egginton, J., Van Ness, B., Van Ness, R., 2013. Quote Stuffing. Working Paper, Louisiana Tech University.

Farmer, J., Joshi, S., 2002. The price dynamics of common trading strategies. Journal of Economic Behavior and Organization 49, 149–171.

Gai, J., Yao, C., Ye, M., 2013. The Externalities of High-Frequency Trading. Working Paper, University of Illinois.

Gaure, S., 2013. OLS with multiple high dimensional category variables. Computational Statistics & Data Analysis 66, 8–18.

Hagströmer, B., Nordén, L., 2013. The diversity of high-frequency traders. Journal of Financial Markets 16, 741–770.

Hendershott, T., Jones, C., Menkveld, A.J., 2011. Does Algorithmic Trading Improve Liquidity? The Journal of Finance 66, 1–33.

Jarrow, R., Protter, P., 2012. A Dysfunctional Role of High Frequency Traders in Electronic Markets. International Journal of Theoretical and Applied Finance 15, 1–15.

Johnson, N., Zhao, G., Hunsader, E., Qi, H., Johnson, N., Meng, J., Tivnan, B., 2013. Abrupt rise of new machine ecology beyond human response time. Scientific reports 3, 1–7.

Jones, C., 2013. What Do We Know About High-Frequency Trading? Working Paper, Columbia Business School.

Kalev, P., Liu, W.-M., Pham, P., Jarnecic, E., 2004. Public information arrival and volatility of intraday stock returns. Journal of Banking & Finance 28, 1441–1467.

Kim, S., Murphy, D., 2013. The Impact of High-Frequency Trading on Stock Market Liquidity Measures. Working Paper, Georgia Institute of Technology.

Kirilenko, A., Kyle, A., Samadi, M., Tuzun, T., 2011. The flash crash: The impact of high frequency trading on an electronic market. Working Paper, University of Maryland.

McInish, T., Upson, J., 2012. Strategic Liquidity Supply in a Market with Fast and Slow Traders. Working Paper, University of Texas.

Quantitative Services Group, 2010. Liquidity Charge & Price Reversals: Is High Frequency Trading Adding Insult to Injury? Available at www.tradersmagazine.com/news/QSG_Study.php, last accessed 02/07/2014.

Scholtus, M., van Dijk, D., Frijns, B., 2014. Speed, algorithmic trading, and market quality around macroeconomic news announcements. Journal of Banking & Finance 38, 89–105.

Securities and Exchange Commission, 2005. Regulation NMS [Release No. 34-51808; File No. S7-10-04].

Securities and Exchange Commission, 2007. Regulation NMS. Final rule; extension of compliance dates [Release No. 34-55160; File No. S7-10-04].

Securities and Exchange Commission, 2010. Concept Release on Equity Market Structure [Release No. 34-61358; File No. S7-02-10].

Securities and Exchange Commission, 2011. Risk Management Controls for Brokers or Dealers with Market Access [Release No. 34-64748; File No. S7-03-10].

Sowers, R., Kirilenko, A., Meng, X., 2013. A Multiscale Model of High-Frequency Trading. Algorithmic Finance 2, 59–98.

Tabb, L., 2013. Time to Determine the Fate of Reg NMS. Available at http://tabbforum.com/opinions/time-to-determine-the-fate-of-reg-nms, last accessed 02/07/2014

Tong, L., 2013. A Blessing or a Curse? The Impact of High Frequency Trading on Institutional Investors. Working Paper, University of Iowa.

Zawadowski, Á., Andor, G., Kertész, J., 2006. Short-term market reaction after extreme price changes of liquid stocks. Quantitative Finance 6, 283–295.

Zigrand, J., Cliff, D., Hendershott, T., 2011. Financial stability and computer based trading. Foresight, Government Office for Science. The Future of Computer Trading in Financial Markets Paper 1, 6–23.

V. Conclusions and Implications for Research

This chapter summarizes the contributions of the essays to their respective literatures, including new publications, and derives up-to-date implications for further research.

1. Essay 1 – Improving Performance of Corporate Rating Prediction Models by Reducing Financial Ratio Heterogeneity

The contribution of this essay to the literature is twofold. First, the approach to reduce financial ratio heterogeneity fills a methodological gap in cases where a data sample characterized by economically diverse groups is too small to construct separate rating models. Second, by laying out a framework of performance levers – factor definition, factor transformation and the choice of classification algorithm – we have provided a toolkit to future researchers in credit risk modeling.

Since its publication in 2008, the article has been cited by numerous later published studies and working papers. The particular issue we faced – building a multi-industry model in the small customer segment of large corporates – has not received much scholarly attention. In the much larger customer segments of small and mid-sized companies, industry or regional heterogeneity can be addressed more efficiently by constructing different group-specific models (Karas & Režňáková, 2015). The authors also show that different predictors are relevant for different sectors and regions; therefore, separate models are the superior approach.

Most citations refer to our second contribution, the framework of performance levers. Karas & Režňáková (2013), completely follow its steps and also employ Box-Cox transformation to normalize predictor variables. Plus, the field has advanced significantly with respect to *classification algorithms*. At the time of our study, we concluded that there is little potential for further performance increases by using better algorithms. Therefore we focused on improving predictive power by extracting better features from the data and transforming them so that they are statistically well-behaved in the model. Kukuk & Rönnberg (2012) as well as Karas & Režňáková (2014) reference our conclusion and show that new classification algorithms do in fact increase predictive power. Also, increases in available computational power have led to the use of much more complex algorithms such as decision trees (Delen et al., 2013; Olson et al., 2012) and semi-parametric methods (Hwang et al., 2010).

Regarding the other performance levers, factor definition, selection and transformation, there have also been numerous advancements. Several studies (Hernandez Tinoco & Wilson, 2013; Li & Miu, 2010) have combined financial ratio information with macro-economic variables and financial market variables to achieve improved predictive power. Our idea of extracting additional information from multi-year transformations of financial ratios is picked up by Volkov & Van den Poel (2012), who show that sequences of financial ratios increase performance. For variable selection, our study used stepwise regression to find a “core model” of the top four or five most significant predictors, and then proceeded with manual variable selection. Tian et al. (2015) show that advanced algorithms such as LASSO are able to select robust variable combinations without further manual intervention.

In sum, it appears safe to state that the best credit risk models will combine advanced classification algorithms with multiple sources of predictive factors – beyond accounting data and financial ratios, and explicitly addressing country and industry level effects – and factor transformations to achieve optimal performance.

2. Essay 2 – Exploiting Attention-Driven Mispricing: Evidence from Actual-Dollar Trading

This essay contributes to three literatures in financial economics: tests of *market efficiency*, studies on *recommendation* and *attention effects* on asset prices, and finally, on the *limits of arbitrage*.

In the literature on *market efficiency*, this study appears to be the first performing an out-of-sample test with real money. Some recent studies come close to this validation approach by observing *other* traders investing real money (e.g. Seasholes & Wu, 2007). Furthermore, instead of relying on *in-sample* results, recent studies take the realistic perspective of a real-time optimizing investor, by evaluating the *out-of-sample* performance of trading strategies (Fang et al., 2014) and asset pricing models (Lewellen, 2015) in a walk-forward test. The essay's second contribution is the finding of an economically small but statistically significant *market inefficiency*. To constitute a large-scale deviation from market efficiency, future research would have to show that the phenomena driving the strategy's profits – (1) attention-driven mispricing (Engelberg et al., 2012) and (2) the observed intraday dynamics during its reversal (first shown in this essay) – are pervasive.

Recent studies such as Yuan (2015) and (Zhang & Wang, 2015) indicate that *attention effects* apply universally. This makes sense given that real-life investors suffer from bounded rationality: they cannot process unlimited amounts of information and therefore have to focus their mental capabilities on stocks that pique their interest. Attention determines where investors allocate their limited research time to acquire firm-specific information (Dong & Ni, 2014). Attention translates into different speeds of how the market prices new information (Drake et al., 2015). Yuan

(2015) shows that market-wide attention effects have market-wide impact on returns, whereas Zhang & Wang (2015) show that attention-driven mispricings and subsequent reversals are prevalent in the Chinese stock market. Attention effects are not limited to the stock market, they occur in the forex markets as well (Goddard et al., 2015). To this literature on attention, and as a subset of that, the literature on *recommendation effects*, the essay contributes two findings. First, it answers the question raised by Engelberg et al. (2012) by showing that Cramer events are actually exploitable. Second, while previous studies suggest a monotonous correction of the initial mispricing, we find that *mispricing re-widens* intraday.

These complex intraday price patterns support findings from theoretical studies on *limits to arbitrage.* To correct a mispricing, arbitrageurs have to coordinate their efforts (Abreu & Brunnermeier, 2002). For a rational arbitrageur, who anticipates that short-term noise traders overwhelm arbitrageurs and exacerbate a mispricing, engaging in this destabilizing speculation is a rational strategy (De Long et al., 1990). Piccione & Spiegler (2014) extend this model and argue that forward-looking arbitrageurs can both exacerbate mispricing and reduce it, depending on the circumstances. In reality, instead of one archetypical rational arbitrageur, there are different levels of arbitrageur sophistication (Milian, 2015). The live trading results in this essay support this, as part of its profits stem from timing the market better than an assumed "average arbitrageur". Thus, its profits are "paid" not only by irrational retail investors, but also by suboptimal strategies of less sophisticated arbitrageurs.

This suggests two questions for future research. First, do other mispricings re-widen intraday? And second, can we show – ideally with actual trading data from brokers – that arbitrageurs coordinate and compete for profits by timing their actions?

3. Essay 3 – High Frequency Trading Intensifies Intraday Extreme Events in Stock Returns

This study adds empirical evidence to the ongoing debate between researchers, regulators and market participants on the role and impact of HFT. Market quality is the focus of this debate, which examines the impact of HFT activity on measures of market efficiency and price discovery, trading costs and liquidity, and volatility and extreme events in stock prices. This essay focuses on the latter and makes two contributions: First, introducing a measure of intraday extreme events which captures unusually large price moves in periods of up to one hour. Second, showing empirically that HFT causes and contributes to these large price changes. This lends empirical support to theoretical models and reports by market participants suggesting numerous mechanisms by which HFT could cause intraday extreme events.

Recent published and working papers contribute to this debate on whether HFT dampen or contribute to large short-term price moves. Brogaard et al. (2015). find that during the top 0.01% largest 10-second price moves high-frequency traders act as net liquidity suppliers, while non-high-frequency traders act as net liquidity demanders. Moreover, high-frequency traders are active liquidity providers during price jumps that result in permanent price changes, absorbing the most informed order flow. This implies that in the very short-term, HFT market making strategies might be dominant and therefore reduce volatility. However, this does not necessarily contradict our findings. First, our research question is on the effect of HFT on intraday extreme events of up to *one hour* in duration, i.e. a timeframe 600 times as long. Van Kervel & Menkveld (2015) show that time makes a difference. High frequency traders lean against institutional orders initially, which dampens very short-term volatility. Yet for longer-lasting institutional orders, HFTs jump the bandwagon and trade in the direction of the order, competing for liquidity, and exacerbating its price impact. Second, results also differ by the source of the data sample. Caivano (2015) investigates the trading activity of 14 identified HFT firms and investment banks with significant HFT activity in large Italian stocks and shows that they cause volatility increases for measurement periods from 10 seconds to 10 minutes. The authors also point out that the method of identifying HFT activity can impact results. If one particular HFT strategy dominates the activity measure, identified effects of HFT will reflect this potential bias. Both the U.S. SEC (2014) and (Foucault & Biais, 2014) point out that the sample used by Brogaard et al. (2015) might be biased towards volatility dampening HFT market making strategies.

The issue of welfare effects from HFT and the appropriate response via regulation and alternative market designs is gaining traction in the literature. The current environment leads to a suboptimal equilibrium in which HFT firms engage in a costly technological arms race for speed, resulting in overinvestment in technology which does not create any societal benefit (Biais et al., 2015). In her review of HFT market microstructure, O'Hara (2015) calls for changes that restore fairness in trading between HFTs and non-HFTs. Recent studies propose alternative market designs that would restore fairness and end the technological arms race: First, by simply introducing a delay of a few milliseconds for all orders (Baldauf & Mollner, 2015), and second, by moving towards frequent batch auctions several times per second rather than trading in continuous time (Budish et al., 2015).

To determine the impact of HFT unequivocally and conclude the HFT debate, future research should full use order book data and a sample of actual orders by a representative set of HFT firms which covers all different forms of HFT strategies.

References (Chapters I+V)

Abreu, D., Brunnermeier, M. (2002). Synchronization risk and delayed arbitrage. Journal of Financial Economics 66, 341–360.

Baldauf, M., Mollner, J. (2015). High-Frequency Trading and Market Performance. Working Paper, University of British Columbia.

Barinov, A. (2015). Why does higher variability of trading activity predict lower expected returns? Journal of Banking & Finance 58, 457–470.

Biais, B., Foucault, T., Moinas, S. (2015). Equilibrium fast trading. Journal of Financial Economics 116, 292–313.

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A., Sokolov, K. (2015). High-Frequency Trading and Extreme Price Movements. Working Paper, University of Washington.

Budish, E., Cramton, P., Shim, J., Athey, S., Ausubel, L., Aze-Vedo, E., Yang, R. (2015). The High-Frequency Trading Arms Race: Frequent Batch Auctions As a Market Design Response. The Quarterly Journal of Economics, 130, 1547–1621.

Caivano, V. (2015). The Impact of High-Frequency Trading on Volatility. Evidence from the Italian Market. Working Paper, Italian Securities and Exchange Commission (CONSOB).

Chacko, G., McMillan, R., Neumar, K. (2014). Optimal Diversification. Journal of Investment Strategies, 3, 59–94.

Chakrabarty, B., Jain, P., Shkilko, A., Sokolov, K. (2014). Speed of market access and market quality: Evidence from the SEC naked access ban. Working Paper, Saint Louis University.

Chipperfield, J. (2013). What Makes a “Good” Model? Retrieved September 13, 2015, from https://josephchipperfield.wordpress.com/2013/09/27/what-makes-a-good-model/

Conrad, J., Wahal, S., Xiang, J. (2015). High-frequency quoting, trading, and the efficiency of prices. Journal of Financial Economics 116, 271–291.

De Long, J., Shleifer, A., Summers, L., Waldmann, R. (1990). Positive feedback investment strategies and destabilizing rational speculation. Journal of Finance 45, 379–395.

Delen, D., Kuzey, C., Uyar, A. (2013). Measuring firm performance using financial ratios: A decision tree approach. Expert Systems with Applications 40, 3970–3983.

Dong, Y., Ni, C. (2014). Does Limited Attention Constrain Investors' Acquisition of Firm-specific Information? Journal of Business Finance & Accounting 41, 1361–1392.

Drake, M., Gee, K., Thornock, J. (2015). March Market Madness: The Impact of Value-Irrelevant Events on the Market Pricing of Earnings News. Contemporary Accounting Research (forthcoming)

Engelberg, J., Sasseville, C., & Williams, J. (2012). Market Madness? The Case of Mad Money. Management Science 58, 351–364.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance 25, 383–417.

Fama, E., French, K. (1992). The Cross-Section of Expected Stock Returns. The Journal of Finance 47, 427–465.

Fama, E., Macbeth, J. (1973). Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy 81, 607–636.

Fang, J., Jacobsen, B., Qin, Y. (2014). Predictability of the simple technical trading rules: An out-of-sample test. Review of Financial Economics 23, 30–45.

Foucault, T., Biais, B. (2014). HFT and Market Quality. Bankers, Markets and Investors 128, 5–19.

Goddard, J., Kita, A., Wang, Q. (2015). Investor attention and FX market volatility. Journal of International Financial Markets, Institutions and Money 38, 79–96.

Hansen, P. (2005). A Test for Superior Predictive Ability. Journal of Business & Economic Statistics 23, 365–380.

Hernandez Tinoco, M., Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. International Review of Financial Analysis 30, 394–419.

Hwang, R.-C., Chung, H., Chu, C. (2010). Predicting issuer credit ratings using a semiparametric method. Journal of Empirical Finance 17, 120–137.

Karas, M., Režňáková, M. (2013). Bankruptcy Prediction Model of Industrial Enterprises in the Czech Republic. International Journal of Mathematical Models and Methods in Applied Sciences 7, 519–531.

Karas, M., Režňáková, M. (2014). To What Degree is the Accuracy of a Bankruptcy Prediction Model Affected by the Environment? The Case of the Baltic States and the Czech Republic. Procedia - Social and Behavioral Sciences 156, 564–568.

Karas, M., Režňáková, M. (2015). Are Bankruptcy Predictors specific for a given Time or Branch? In Proceedings of the 7th International Scientific Conference - Finance and Performance of Firms in Science, Education and Practice, pp. 571–583.

Khovansky, S., Zhylyevskyy, O. (2013). Impact of idiosyncratic volatility on stock returns: A cross-sectional study. Journal of Banking & Finance 37, 3064–3075.

Kukuk, M., Rönnberg, M. (2012). Corporate credit default models: a mixed logit approach. Review of Quantitative Finance and Accounting 40, 467–483.

Kuntz, K., Sainfort, F., Butler, M., Taylor, B., Kulasingam, S., Gregory, S., Kane, R. L. (2013). Decision and Simulation Modeling in Systematic Reviews. Methods Research Report. Rockville, Maryland.

Lewellen, J. (2015). The Cross-section of Expected Stock Returns. Critical Finance Review 4, 1–44.

Li, M.-Y., Miu, P. (2010). A hybrid bankruptcy prediction model with dynamic loadings on accounting-ratio-based and market-based information: A binary quantile regression approach. Journal of Empirical Finance, 17, 818–833.

Malkiel, B. (2003). The Efficient Market Hypothesis and its Critics. The Journal of Economic Perspectives 17, 59–82.

Marshall, B., Cahan, R., Cahan, J. (2008). Does intraday technical analysis in the U.S. equity market have value? Journal of Empirical Finance, 15, 199–210.

Milian, J. (2015). Unsophisticated Arbitrageurs and Market Efficiency: Overreacting to a History of Underreaction? Journal of Accounting Research 53, 175–220.

O'Hara, M. (2015). High frequency market microstructure. Journal of Financial Economics 116, 257–270.

Olson, D., Delen, D., Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. Decision Support Systems, 52, 464–473.

Park, C., Irwin, S. (2007). What Do We Know about the Profitability of Technical Analysis? Journal of Economic Surveys, 21, 786–826.

Piccione, M., Spiegler, R. (2014). Manipulating market sentiment. Economics Letters 122, 370–373.

Režňáková, M., Karas, M. (2014). Identifying bankruptcy prediction factors in various environments: A contribution to the discussion on the transferability of bankruptcy models. International Journal of Mathematical Models and Methods in Applied Sciences 8, 69–74.

Seasholes, M., Wu, G. (2007). Predictable behavior, profits, and attention. Journal of Empirical Finance 14, 590–610.

SEC. (2014). Equity Market Structure Literature Review Part II: High Frequency Trading. U.S. Securities and Exchange Commission, 1–37.

Tian, S., Yu, Y., Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. Journal of Banking & Finance 52, 89–100.

Todd, P., Gigerenzer, G. (2000). Précis of Simple heuristics that make us smart. Behavioral and Brain Sciences 23, 727–741.

Van Kervel, V., Menkveld, A. (2015). High-Frequency Trading around Large Institutional Orders. Working Paper, Pontificia Universidad Católica de Chile.

Volkov, A., & Van den Poel, D. (2012). Extracting Information from Sequences of Financial Ratios with Markov for Discrimination: An Application to Bankruptcy Prediction. In 2012 IEEE 12th International Conference on Data Mining Workshops, pp. 340–343

White, H. (2000). A Reality Check for Data Snooping. Econometrica 68, 1097–1126.

Yuan, Y. (2015). Market-wide attention, trading, and stock returns. Journal of Financial Economics 116, 548–564.

Zhang, B., Wang, Y. (2015). Limited attention of individual investors and stock performance: Evidence from the ChiNext market. Economic Modelling 50, 94–104.