Technische Universität München
TUM School of Management

Microstructure of Energy Markets: Theory and Empirics

Vadim Gorski

Vollständiger Abdruck der von der TUM School of Management der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Volkswirtschaftslehre (Dr. oec. publ.) genehmigten Dissertation.

Vorsitzender: Prof. Dr. Christoph Ungemach, Ph.D.
Prüfer der Dissertation: 1) Prof. Dr. Sebastian Schwenen
2) Prof. Dr. David Wozabal

Die Dissertation wurde am 21.05.2021 bei der Technischen Universität München eingereicht und durch die TUM School of Management am 15.11.2021 angenommen.

Acknowledgements

First, I would like to thank my supervisor, Sebastian Schwenen for his ongoing support, guidance and the countless hours of discussions during the last four years. Thank you for supporting me in my explorations and giving me an opportunity to follow my interests.

I also want to thank David Wozabal for his willingness to serve as my second reviewer. A special thanks goes to Moritz for being a great friend and colleague. I would also like to thank my friends for their outstanding support during this time. Thank you Vlady, Martin, David, Nico, Marvin, Annette, Karina.

Last, I want to thank my family and Nicola.

Abstract

Energy markets worldwide are facing a transformation amid rising energy needs and climate change mitigation efforts. One part of this transformation consists in the move towards liberalized markets where power can be traded like any other commodity. This change calls for new approaches to market design, and poses various challenges for regulators, exchanges, and market participants due to the peculiarities of power as a commodity.

On the other hand, the rising prevalence of computerized markets and the availability of high-frequency market data present a unique opportunity to study the markets and develop policies. In this thesis, I explore these two aspects of modern energy markets.

First, I show how demand elasticity can be estimated from high-frequency market data using renewable energy feed-in as an exogenous explanatory variable. My empirical findings show that demand elasticity can be approximated precisely using the proposed approach. Second, I develop a theoretical market model of continuously trading markets with workup sessions. I show how workup sessions can mitigate bid shading and foster more effective reallocation of financial assets under limited liquidity. I present an empirical case study based on the Nasdaq Nordic power market. Third, I show how the regulation of insiders' trading impacts the excess returns achieved by corporate insiders. In particular, my research reveals that regulators face a trade-off between fairness in the market and increased welfare by the signaling effect of insiders' trades. I derive disclosure policy recommendations based on my findings.

My findings contribute towards a theoretical and empirical understanding of markets with implications for power market design.

Contents

List of Figures

List of Tables

1 Introduction

This introduction serves as an overview of the most recent developments in energy market microstructure and how these motivate my research. First, I present a summary of how the power markets have developed in recent years with a focus on the computerization of trading and market mechanisms. Second, I lay out in more detail how these changes affected market mechanisms in power trading. I proceed to discuss the possibilities of widely available financial market data and how it can be used in econometric research. I then provide an outlook on the key challenges in designing power markets going forward. Next, I present an overview of my research, my main contributions, and what these imply for the power market design of the future. Last, I present the structure of my thesis.

1.1 Developments in Power Market Microstructure

The wave of power market liberalization initiated in the 90s brought upon a fundamental transformation of the way power is traded and distributed. The changes in the power market microstructure triggered by this development are still ongoing.

Regulators and exchanges are addressing a myriad of aspects of new power markets when designing their policies. On the one hand, market design is supposed to foster competition and, therefore, more efficient market outcomes (Wilson, 2002). On the other hand, it has to address various topics arising from the shift towards intermittent power sources such as investment incentives, renewable support schemes, the complexity of grid operations, and the increasingly short-term nature of demand and supply shifts. Addressing all of these topics raises multiple questions in market design.

1.1.1 Trading Mechanism Design

The choice of trading mechanisms to exchange power between traders among markets with different maturities is one of such questions.

Electricity is traded on multiple markets with varying maturities due to its special properties. In particular, there are no economically viable ways to store it at scale. In addition, supply and demand need to match exactly at any given point in time, and prices can be subject to high volatility (Eydeland and Wolyniec, 2002). To account for these properties, regulators and exchanges respond by introducing an increasingly complex set of markets to trade electricity on. The two most prevalent choices for market design in this area are various types of double auctions and continuous limit order books (LOBs).

Power markets and their mechanisms differ depending on the maturity of the underlying financial contracts traded on these markets and depending on the market participants trading them. Today, exchanges are offering multiple products: from long-term ones going multiple years into the future to very short-term products (sub-hourly). At the same time, market participants active in these markets can have very different natural positions, risk preferences, and planning horizons. For example, power producers are typically taking a net long position with long-term planning horizons, as their cashflow can be more dependent on price fluctuations than the cashflow of an industrial participant for whom electricity is just one input factor. A retailer can hold a long or short position with a medium planning horizon. Last, an end consumer is typically short and plans the position not too long in advance in order to profit from possible short-term price fluctuations.

On the long-term end of the spectrum, there are markets with a planning horizon spanning up to 10 years, e.g., in the form of power purchase agreements (PPAs) which can be used to reduce the risk for renewable power procurement and investment (Hundt et al., 2020). PPAs represent a contract to exchange power at a future date for a fixed price and are usually exchanged over-the-counter (OTC).

There are medium-term markets in the form of classical futures markets. Here, participants exchange electricity with a planning horizon of up to 3 years. These markets are mostly organized as LOBs. Large participants also exchange forward contracts OTC. In terms of total volume traded, futures markets generally represent the most significant ones, especially in Germany.

They serve a multitude of roles, including hedging risks, optimal price discovery, mitigation of market power exertion (Allaz and Vila, 1993, Mansur, 2007) and as a tool to price investments (Ausubel and Cramton, 2010). As such, their design plays a key role. Examples of issues addressed by market design in this area are front-running, bid shading (and more generally strategic bidding), market power exertion, appropriate demand and supply response, and planning uncertainty. Planning uncertainty can arise because the price forward curve deviates a lot from the actual spot price at a future date. Some of these problems, such as strategic order placement (Budish et al., 2015) or market power exertion (Bushnell et al., 2008) can be addressed by appropriate changes in market mechanisms.

On the short-term side, exchanges are offering day-ahead products and intraday products, allowing traders to exchange electricity only hours before delivery, a counterpart to spot trading in other commodities. These markets are partly organized as double auctions and partly as continuous LOBs. They are also split with regard to delivery periods. Exchanges are offering an ever-rising number of products batching multiple hours to best approximate the needs of suppliers and consumers. For the European Energy Exchange (EEX), these products span, e.g., base (electricity delivered in all hours of the day, 7 days a week), peak (hours 8 to 20, Monday to Friday), off-peak, and many others. At the same time, the amount of sub-hourly products is rising, too. With an increasing share of renewable power, these markets have seen steady growth in terms of volume, as the need to rebalance the position and respond to the newest information is on the rise. This, in turn, has led to multiple discussions on optimal mechanisms, as, e.g., in Neuhoff et al. (2016), Ocker and Jaenisch (2020).

Last, to balance the very short-term mismatch of demand and supply, many countries have introduced balancing markets. Balancing power is typically provided at a high premium and can be dispatched within seconds to minutes. In these markets, participants are remunerated for both, provision of flexibility and actually delivering electricity when required. In these markets, regulators are also facing design questions, in particular, due to the tight connection between the physical delivery and the financial transactions behind it. In addition, these markets play a crucial role in network stability and present a strategically important part of the electricity infrastructure (van der Veen and Hakvoort, 2016).

The high impact of these developments is not limited to exchanges. Similar questions also arise on the other side of the screen, where participants active in these markets are trying to find optimal strategies to trade or hedge power. Continuous LOBs pose high requirements on the trading infrastructure, the decision process of traders, and the effective execution of the trades themselves. A high need for coordination between the long-term and the short-term markets makes these requirements even more complex. Modeling such decision processes (Lai, 2018) is just the first part of the problem of finding a suitable trading strategy. Given that this problem is solved effectively, optimal liquidation and order placement are another crucial part of trading in LOB markets (Obizhaeva and Wang, 2013, Cont and Kukanov, 2017). Optimal liquidation is closely related to optimal algorithmic trading, which is continuously gaining importance as power markets become more and more automated and the need for short-term rebalancing increases (Wang and Yu, 2019). To sum up, the question of optimal trading mechanisms is prevalent across all power markets. The interplay between markets of different maturities makes the question of optimal design even more complex. The goal of an optimal market design is to allow market participants to exchange power according to their specific risk preference and volumetric requirements while avoiding the exertion of market power and strategic behavior. This entails optimal price and size discovery.

1.1.2 Increasing Availability of High-frequency Data

Computerized markets are naturally generating large amounts of high-frequency data. For a typical LOB market, the data is often quoted on a millisecond scale, as this is the typical scale algorithmic trading takes place on (Goldstein et al., 2014). While this development started in the 90s for stock, fixed income, or derivatives markets, algorithmic trading in power markets has gained interest only recently (Baltaoglu et al., 2018, Wang and Yu, 2019).

This development poses a second key challenge for regulators and market participants. For regulators, an analysis of this data presents an integral part of measuring the efficacy of mechanisms, incentive schemes, or regulations. For market participants, processing market data poses an important prerequisite for decision making and serves as an input to, e.g., trading algorithms.

One important application of market data analysis is the measurement of demand elasticity. Demand elasticity in power markets is connected to multiple aspects of their organization, including demand flexibility and demand response, real-time pricing, strategic behavior, and integration of renewable power sources (Borenstein, 2005, Bompard et al., 2007, Clastres and Khalfallah, 2015, Gold et al., 2020). A more recent development towards smart grids and more localized markets is another key application of demand elasticity, as price response in such markets is an important indicator of their allocative efficiency (Fabra et al., 2021).

For policy considerations, the requirement on elasticity estimation models is even higher, as policy-makers are typically interested in causal relationships between price and quantity demanded as opposed to the simple association between the two.

For market participants, demand elasticity serves as an important input to a variety of models. For example, producers employ elasticity estimates for constructing optimal day-ahead bidding strategies (Hajati et al., 2011). Another application are fundamental market models used, e.g., for investment decisions into new technologies or PPA pricing (Gerbaulet and Lorenz, 2017).

The increasing availability of high-frequency data in computerized markets spurred a new strain of research into statistical methods to analyze it. In finance, machine learning has seen a steady increase in popularity with a wide range of applications in the pricing of derivatives, forecasting, algorithmic trading, and many others. As methods from finance and industrial organization become increasingly blended (Kastl, 2017), these methods find their way into the industrial organization, too. Recent studies show that these methods are widely applied in energy economics and finance (GHO, 2019). Despite the advantages of machine learning applied to high-frequency data, researchers and practitioners are also facing some risks, as the properties of these methods are not always well-understood (Abadie and Kasy, 2019).

As a consequence, there are promising research avenues for using high-frequency data and studying machine learning approaches in the context of market design and market microstructure.

1.2 Power Trading Mechanisms

In liberalized power markets, trading represents a key mechanism to allocate goods. It can take place in various forms, including double auctions, price-driven markets or continuous limit order books (order-driven markets), known as lit markets. In addition, some exchanges offer dark pool trading. These venues offer participants to submit large orders without making these public. This process can take place with a wide variety of mechanisms. In addition, bilateral trading and OTC trading are also prevalent in today's markets.

Researchers, regulators, and market participants employ various metrics to judge these trading mechanisms. While an overview of desirable qualities of trading mechanisms is a very far-reaching topic, there are some standard properties that are commonly discussed by both, academics and practitioners. These properties include the speed and precision of price and size discovery, reallocation speed, proneness to strategic behavior, pre-trade and post-trade transparency.

LOBs present the most frequently used mechanism among all of today's biggest exchanges. As of 2020, the market capitalization of exchanges employing some form of LOBs amounted to 82% of the worldwide market capitalization of all exchanges (World Federation of Exchanges, 2021). For this reason, LOBs have seen a steadily growing interest in the literature.

The current understanding of LOBs is fragmented, as there are different perspectives on viewing these markets. From an economic perspective, researchers are dealing with questions such as welfare, strategic behavior, and market power. From an operations research perspective, optimal liquidation strategies are an important topic. The increasingly low-latency nature of these markets has also led to substantial research in algorithms and automated trading logic.

The question of whether continuous LOBs can be further improved by augmentation with additional mechanisms has seen a lot of interest, too. Empirically, it is known that dark pools present an important additional trading venue, especially in equity markets and fixed income markets, where they amount to around 50% of total traded volume (Fleming et al., 2018). Yet, their overall net impact is ambiguous. On the one hand, dark pools are known to better serve the needs of large institutional traders, as they limit strategic front-running and usage of iceberg orders (Frey and Sandås, 2017). On the other hand, dark pools were found to decrease the liquidity in the lit market (Nimalendran and Ray, 2014, Johann et al., 2019) and thereby decrease their effectiveness of price discovery. The net benefit is therefore not clear and has led to controversial discussions in the literature. Researchers struggle to find a definitive answer here, as these markets are difficult to model theoretically and commonly do not offer their data for academic analysis. The question of their role for market efficiency, therefore, remains open and requires efforts from theoretical and empirical researchers alike.

1.3 Econometrics and Machine Learning

Turning to the use of machine learning for analyzing financial market data, recent years have seen a variety of applications combining it with econometric methods (Shmueli and Koppius, 2011). While the main focus of econometrics is often on structuring the models and causal inference, machine learning methods have the advantage of approaching problems in a data-driven way without putting much structure on them.

A combination of these methods is found in multiple applications. The first prominent example is a data-driven way to select parameters in econometric models when no further information about these parameters is available. Machine learning techniques also find application in modeling non-linear relationships in econometric models (Varian, 2014).

Borrowing variable selection tools from machine learning to build econometric models is another important example. Here, cross-validation and regularization present two well-known approaches. Cross-validation allows improving the out-of-sample fit by artificially splitting up the dataset into parts and making sure that the estimated model is a good fit on each part (fold). Regularization has a similar property in that it artificially reduces the in-sample fit in the hope of improving the out-of-sample fit.

One important application for variable selection on the econometrics side are instrument variable (IV) models. These used to be estimated based on economic reasoning for the choice of variables. Even though this economic reasoning is still important for the preselection of instruments, regularization, and cross-validation can help in selecting the right set of instruments. While there are multiple prior works on the efficiency of such techniques, field tests with known benchmarks are rare. This opens up a possibility for further research.

1.4 Contribution

In my thesis, I contribute to a deeper theoretical and empirical understanding of mechanisms behind today's financial markets in terms of price and size discovery, with a focus on electricity markets. My findings are relevant to both, regulators and market participants. My thesis consists of three main essays, each with a distinct research question and contribution.

First, I show how regularization can be used to better estimate demand elasticities from high-frequency market data. These estimates can, in turn, be used to judge the effectiveness of policies that aim to flexibilize the demand side in power markets. Recent works have shown how demand elasticity can be used as a measure of the effectiveness of real-time pricing (Fabra et al., 2021). Second, I show how augmenting classic LOB markets with mechanisms such as workup sessions (Duffie and Zhu, 2017) can alter the behavior of market participants and influence their strategies. For regulators, these findings imply possible changes to trading mechanisms that can ensure faster reallocation of financial contracts. For market participants, I explain how an optimal liquidation strategy can be derived theoretically. I also demonstrate empirically that the optimal liquidation strategy achieves a lower overall cost of liquidity, yielding higher average prices for net sellers. In addition, I lay out how market participants can leverage power market data to optimize their market position and maximize profits. Last, I look into the regulatory aspect of designing efficient financial markets. I show how disclosure of insiders' trading influences market outcomes in terms of fairness and information asymmetry. I build on extant literature on event studies (Corrado, 2011) and show how regulators can use market data to derive optimal policies.

The empirical part of my contributions is based on power market data from the German power market (EEX) and the Nordic power market (Nasdaq Nordic). Both datasets provide detailed data on market participants' behavior, in the case of Nasdaq on the order-book level. I also employ data on insiders' dealings in German, UK, and US-American markets.

In the following, I summarize each of the three essays and lay out their distinct contributions.[1]

1.4.1 Demand Elasticity Estimation with Many Instruments

In the first essay of this thesis, I investigate different methods for demand elasticity estimation in power markets. This estimation is subject to classic identification issues, as the supply and the demand shifts cannot be disentangled based on prices and quantities only. Some classic approaches are known to solve this problem, with the most prominent one being instrument variable estimation (Angrist and Krueger, 2001). Other approaches include cointegration models, structural time series models, and Kalman filter models (Inglesi-Lotz, 2011, Arisoy and Ozturk, 2014).

Researchers typically utilize data on equilibrium (market-clearing) prices and quantities to estimate demand elasticity, mainly due to the unavailability of more precise data reflecting quantities demanded at different (non-market-clearing) prices.

This is problematic for several reasons. First, there is no way to judge the quality of estimates stemming from equilibrium price-quantity data. Researchers are employing methods such as

[1] I use the first person throughout, even though the essays are based on joint work with co-authors

instrument variable regression to overcome the simultaneity bias, but these methods do not guarantee precise estimates (Angrist et al., 2000). In fact, the instrument variable (IV) estimates can be even further away from the true estimate when weak or wrong instruments are selected (Jiang, 2017). Second, without any external way to judge the quality of the estimates, research on the methods of estimation themselves is difficult. While some properties of such estimators can be shown in theory and in simulation studies, as in, e.g., Gold et al. (2020), this is often an imprecise approximation to what would happen with real-world data. Theoretical properties often only hold asymptotically. Simulation studies mostly assume a certain parametrization, and their robustness with respect to different parametrizations can be questionable.

I make use of a unique setting in terms of data availability to address this question. I employ data from the German power markets where both, the equilibrium price-quantity pairs and the submitted demand curves (quantity demanded for different levels of price) are directly observable. The former allows me to estimate demand elasticity using instrument variables, as regularly applied by empirical researchers. The latter allows me to compute the actual underlying demand elasticity and to compare my estimates to it. The actual demand elasticity, therefore, serves as a benchmark to judge the performance of different estimation methods.

Another important ingredient for elasticity estimation models is given by strong and preferably exogenous instruments, in my case by supply shifters. For supply shifters on the German power market, I, therefore, make use of the predicted in-feed from renewable power sources. This data is publicly available, exogenous, and a strong supply shifter, as the total installed capacity is constantly expanding. I also utilize fossil fuel prices from day-ahead markets and CO_2 prices as instruments.

In terms of estimation methods, I vary between classic IV models with controls and regularized IV models. I propose regularization in the "first stage" of the estimation as a way to select the most suitable instruments (Belloni et al., 2012). I investigate how LASSO, RIDGE, and Post-LASSO type of estimators perform in this instrument selection problem.

I show empirically that, especially in cases with possibly unfit instruments, whose local average treatment effect (LATE) deviates from the effect in the entire population, regularization can significantly improve the resulting estimates. This observation is robust with respect to the estimation period, controls, and over time.

My results contribute to the literature in two ways. First, I provide precise true estimates of the demand elasticity in the German power market, which can be used in a variety of models building upon the elasticity of demand. Second, I provide reasoning why instrument selection is an important step in IV estimation and proof that the right instrument selection can influence results significantly. This makes the results relevant to both, researchers in industrial organization and practitioners active in power markets.

1.4.2 Optimal Liquidation in Continuous Markets with Workups

In the second essay of the thesis, I look into the problem of optimal liquidation in continuous LOB markets with workup sessions, as defined by Duffie and Zhu (2017). In a workup session, a price for a financial contract is frozen based on the latest market price or another price-determining mechanism. Participants are then allowed to submit bids in terms of quantity only. The process of quantity submissions takes place until one side (either bid or ask) drops out of the session. This mechanism is therefore sacrificing complete market clearing and is designed to enhance the size discovery process. Size discovery sessions are found across all financial markets, including equities, fixed income, and commodities. This kind of mechanism is highly relevant for electricity markets, where large producers and consumers are trading their portfolios far in advance.

The motivation for augmenting continuous LOB with size discovery sessions is straightforward: market participants are typically shading their true bids in fear of possible price impact, especially when trying to liquidate a non-trivial position. This means that they either do not submit their true bids or use techniques such as iceberg orders to trade (Frey and Sandås, 2017).

By this reasoning, a workup session with a frozen price gives market participants an incentive to disclose their desired quantity faster. This, in turn, can lead to a faster liquidation. A faster reallocation speed is a desirable property and is considered welfare-enhancing. For an exchange, it can be a competitive advantage to offer workups, as exchanges compete against each other in terms of trading speed (Pagnotta and Philippon, 2018).

Despite this reasoning, the literature on the role of workup sessions in LOB markets is quite ambiguous. Researchers found that this mechanism can have negative consequences, as it draws away liquidity from the main trading venue, thus decreasing the market's price discovery capabilities. It is also prone to strategic behavior, as market participants can strategically wait for the next workup session before submitting any bids. In addition, such a mechanism is prone to front-running, as the information on which side drops out first is public. This information can then be used in the subsequent continuous market to profit from the temporal imbalance.

There is also some evidence that size discovery sessions can be welfare-enhancing. Theoretical researchers are trying to model the market and the behavior of participants to underpin this point. Empirical works are typically looking at liquidity, bid-ask spreads, reallocation speeds, and other metrics to judge the effectiveness of this mechanism. My essay aims to provide further theoretical and empirical evidence on the role of workups in continuous LOBs.

In the theoretical part of my work, I build upon the continuous-time market model of Almgren and Chriss (2001) and extend it to include limit orders and workup session orders. This allows me to reflect the decision process faced by a trader who is active on an exchange that offers standard LOB trading with market and limit orders but also offers size discovery sessions. Similar to Antill and Duffie (2020), I randomize the arrival of orders in the size discovery session. This allows me to circumvent the problem of traders strategically waiting for the arrival of the next workup session. I also introduce waiting costs to incentivize traders to liquidate their position at the earliest time possible. I find an optimal closed-form solution to this problem by formulating it as a dynamic optimization program and solving the corresponding Hamilton-Jacobi-Bellman equation. My results show that a workup session influences traders' behavior in terms of which orders they place. I also argue that there is an optimal split between the regular, continuous market and the workup market, ensuring the fastest reallocation speed. These results have strong implications for exchanges and regulators aiming to improve the classic LOB trading mechanism.

In the empirical part of my work, I make use of order book level data from the Nasdaq Nordic exchange. I simulate a setting that allows traders to participate in a workup session. Building upon the classic benchmarks for optimal liquidation, such as time-weighted average price (TWAP) or volume-weighted average price (VWAP), I show that markets with workups provide better average prices for both market sides while increasing the reallocation speed. I also demonstrate that the optimal strategy determined in the theoretical part of my work can easily be implemented in a real-world trading algorithm.

1.4.3 Fairness vs. Welfare: Disclosure in Financial Markets

In the third essay of my thesis, I look at the role of regulation, in particular disclosure, in designing efficient financial markets. Company insiders trading companies' stock are obliged to disclose their position. This has several reasons. First, it is assumed that insiders possess superior information about the state of the company. This creates information asymmetry between insiders and other traders. This information asymmetry can be reduced if insiders' trading intentions are known, thereby increasing welfare (Lenkey, 2014). This is called the indicator effect and represents one of the regulatory aims. The second aim is fairness, measured by excess performance an insider can achieve when trading a company's stock. If the excess return deviates from the normal market return[2], the trade is considered unfair. Various empirical studies (Dardas and Güttler, 2011, King et al., 2015, Hodgson et al., 2020) show that insiders actually achieve excess returns. This motivates the need for regulatory intervention. In their endeavor to set the optimal degree of disclosure, regulators face a trade-off. I find that, while reducing fairness in the market, a looser regulation of insiders' trades entails welfare gains due to the positive signaling effect.

In the essay, I shed light on the implications of insiders' trading on market outcomes by analyzing disclosure policies and transactions data on insiders' deals from the German, the UK, and the US-American markets. I show how the seminal event study methodology (MacKinlay, 1997, Corrado, 2011) can be extended to account for both, pre-event and post-event effects. I also argue that compounding of returns plays a key role in such studies (McLean, 2012) and introduce a return metric to account for it. In addition, I provide a possible explanation for differences in excess returns among analyzed markets and explain the role of regulatory enforcement for the efficacy of disclosure policies.

First, I introduce the necessary modifications to the standard event study methods. One such modification is the timing estimator. This estimator relates pre-event, and post-event effects of an insider's trade, thereby allowing me to judge the timing exhibited by corporate insiders. I also introduce a metric for compounded excess returns.

Next, I analyze the dataset on insiders' transactions and show that in all three financial markets, they achieve excess returns. This effect is particularly pronounced in the USA. I relate my findings to the properties of the transaction, such as volume, the duration between the transaction

[2]Normal market return is derived based on a market index.

and its public disclosure, and the corporate level of the insider.

Based on my findings, I provide reasoning why pre-trade or simultaneous disclosure constitutes a better way to enforce fairness. I also argue that the enforcement of disclosure regulations plays a key role in its efficacy, but exhibits substantial drawbacks, especially in the German market.

I conclude the essay with recommendations for disclosure regulation.

1.5 Structure of the Thesis

The remainder of this thesis is structured as follows. Chapter 2 explores how regularization methods can be applied to estimate demand elasticity and presents applied findings for the German power market. In Chapter 3, I present a continuous-time market model with workup sessions and argue theoretically and empirically how these influence market outcomes. Chapter 4 deals with the regulatory side of designing efficient markets via effective disclosure processes. Last, Chapter 5 concludes with an overview of how my research can be applied to improve the market design of financial markets, and in particular, of power markets.

2 | Estimation of Demand Elasticity with Many Instruments: A Machine Learning Approach

Vadim Gorski, Sebastian Schwenen[3]

In settings with high-dimensional data, endogenous regressors and several instrument candidates for IV estimation, the choice of instruments becomes essential to estimate the effect of interest. We explore how regularization techniques can improve inference for the canonical two-stage IV demand estimation problem. We rely on high-frequency data from double auctions, where we observe submitted demand curves that allow us to infer how well the proposed regularization-based IV methods approximate true elasticity. We find that regularization in the first stage correctly selects the most relevant instruments, significantly improves inference, and reduces the mean squared error. We derive suggestions for researchers using machine learning tools in IV models with many instruments.

[3]Author contributions: This essay is based on a joint paper with Sebastian Schwenen. My contribution was the formulation of the research question, the design of the empirical strategy, large parts of empirical work, and the draft of major parts of the paper

2.1 Introduction

Estimating demand elasticity from equilibrium price-quantity pairs is a classic problem in economics. However, researchers often face difficulties to arrive at good estimates, either because of too much or too little data or due to the necessity to find reasonable instruments. In the worst case, estimates become unreliable and introduce bias in policy-relevant counterfactuals.

Instrumental variable (IV) estimation presents one of the key approaches to estimating demand elasticity. The seminal literature on demand estimation using IV concentrates on cases with scarce data, i.e., only few available instruments (Angrist and Krueger, 2001). Recently, the increasing availability of large datasets with a high frequency of observations has encouraged novel research in empirical economics and applications of IV, with its own set of obstacles in handling large volumes of data.

In case of scarce data, researchers find it hard to identify the true effect of interest because of potential bias in the local average treatment effect. In case of too much data, researchers struggle to find the right instruments that are best suited to reduce bias. As marketplaces become increasingly computerized, market participants generate vast amounts of data with many possible instrumental variable candidates. Paramount examples are marketplaces such as Amazon or Uber, which gain increasing focus for the estimation of demand and consumer surplus (Bajari et al., 2015, Cohen et al., 2016).

In this article, we study the usefulness of machine learning approaches for demand estimation in high dimensional, high-frequency settings. We make use of the seminal two-stage IV model and estimate demand elasticity from equilibrium prices, quantities, and numerous supply shifters. As well known, bad instruments correlated with the error term can possibly worsen the bias as compared to using ordinary least-squares (OLS). In large datasets, the availability of many, possibly unfit, instruments increases the risk of misidentification and over-fitting, again introducing bias.

The statistics and machine learning literature has proposed a host of approaches to guide the variable selection and estimation process for the problem at hand.[4] While these methods are originally designed for prediction, a growing literature explores their usefulness for inference of

[4]See, e.g., Varian (2014) and Athey (2018) for comprehensive reviews and implications for empirical economics.

model parameters and underlying economic models (Belloni et al., 2012, 2014, Fessler and Kasy, 2019).

In our study, we apply machine learning techniques to the seminal IV problem and explore their usefulness for improving inference. In particular, we investigate how first-stage regularization performs in selecting suitable instruments, and how this selection affects the second-stage estimates. Specifically, we focus on LASSO, RIDGE, and Post-LASSO estimation (Hoerl and Kennard, 1970, Tibshirani, 1996) of the first-stage model. We follow an applied perspective on settings with many instruments, in our case supply shifters, and apply different IV model formulations to the canonical demand elasticity estimation problem (e.g, Angrist et al., 2000).

Our data stem from computerized double auctions in wholesale electricity markets. This dataset provides a rich test environment for our study as it entails revealed demand curves submitted to the double auction. Observing demand curves enables us to compute the "true" underlying demand elasticity. This, in turn, allows us to evaluate the different estimates from IV regression based on equilibrium prices, quantities, and instruments. Knowing the true elasticity furthermore allows us to infer how well different IV models with and without machine learning approximate the true elasticity.

Our findings show that regularization in the first-stage reduces bias and mean squared error (MSE) of the estimator. This is, we find that in situations with many instrument candidates for IV estimation, the application of regularization in the first stage dominates standard two-stage models.

As our results show, regularization is especially valuable under certain conditions. First, we find that regularization improves inference when instruments are correlated or exhibit outliers. Second, we find that regularization is effective in addressing the weak instrument problem in that it aids the instrument selection process and avoids over-identification. Our dataset includes a variety of arguably exogenous instruments and allows us to test the instrument selection properties of the regularized first-stage estimator. It is commonly known that selecting the seemingly best instrument requires economic reasoning and an understanding of the underlying mechanism (Angrist and Krueger, 2001) while using several plausible instruments at once can introduce additional bias. In this regard, we find that combining economic reasoning with machine learning based instrument selection presents a data-driven answer to this trade-off.

Our findings relate to several strands of research. First, we contribute to the emerging literature on the estimation of demand and consumer surplus with large data and machine learning tools (Bajari et al., 2015, Cohen et al., 2016). While Bajari et al. (2015) use machine learning to predict the treatment effect of price promotions on demand, our approach instead aims at improving the traditional IV demand estimation procedure. Although we conduct our analysis within the demand estimation framework, we believe it is widely applicable to a broader set of IV models with large datasets.

Second, we relate to the literature on demand estimation, that has been applied to a variety of problems and settings in industrial organization (e.g., Genesove and Mullin, 1998, Berry et al., 1995). More specifically, we relate to the literature focusing on estimating demand elasticity in power markets (Lijesen, 2007, Bönte et al., 2015, Boogen et al., 2017), where estimating elasticity has become of particular interest for evaluating policies to implement smart meters and real-time pricing schemes Fabra et al. (2021). As we show, adding machine learning tools to the traditional two-stage IV design significantly improves estimates and, as such, can also be widely used for studying policy counterfactuals.

Last and more broadly, we add to the growing literature at the intersection between applied economics and econometrics on the use of machine learning for causal inference. In a recent paper, Abadie and Kasy (2019) show that data-driven regularization yields estimates that are close to estimators obtained for optimal levels of regularization. We draw from this idea but apply regularization to select among many instruments within a two-stage IV setting, which has previously been studied in Okui (2011) and Belloni et al. (2012). Similar works develop methods for high-dimensional instrumental variable settings, including (Zou, 2006, Belloni et al., 2014, Carrasco and Tchuente, 2015). These works mostly concentrate on the derivation of asymptotic properties of the estimators. Our focus is on the application of these methods to a practical demand estimation problem and the investigation of how these estimators perform "in the field".

The remainder is organized as follows. Section 2.2 introduces theoretical foundations and notation. Section 2.3 describes the dataset, while section 2.4 outlines our empirical strategy. Section 2.5 presents our findings and results. Section 2.6 concludes.

2.2 Model

This section outlines a brief equilibrium market model. Subsequently, we use the model to illustrate how we measure and estimate demand elasticity. Our setup and notation draws from Angrist et al. (2000).

Consider a set of n competitive markets in which buyers and sellers exchange a perfectly homogeneous good.[5] Each single market $i = 1, \ldots, n$ is subject to supply shocks from k supply shifters that we denote as z_j with $j = 1, ..., k$. We write the set of all supply shifters as $Z = [z_1, \ldots, z_k]$. The market supply function hence depends on the realization of Z and can be written as $S_i(p, Z)$. In equilibrium, realized supply and demand determine a market-clearing price-quantity pair $\{p_i^*(Z), q_i^*(p_i^*(Z))\}$.

We denote the demand function in market i as $Q_i(p)$. Furthermore, we assume that demand is stochastic and has a zero mean random component so that $Q_i(p) = Q_i(p, \xi)$ with $\mathbb{E}[\xi|p] = 0$. The error term ξ_i may contain any perturbations due to unobserved consumer behavior in market i.

As the good is traded over i markets, we measure demand elasticity, denoted by ϵ, as the average elasticity over all n markets, formally:

$$\epsilon = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \mathbb{E}[Q_i(p,\xi)]}{\partial p}\frac{p}{Q_i} = \frac{1}{n}\sum_{i=1}^{n} \epsilon_i, \tag{2.1}$$

where ϵ_i is the elasticity in market i. The average demand elasticity as defined in Equation (2.1) is the object of interest that we seek to estimate.

If the realization of the demand curve in market i is observable, say as submitted demand curve in a double auction, we can estimate (2.1) from data on Q_i and p.

Typically, researchers however do not observe realized demand curves. In this case, empirical specifications have to rely on observed equilibrium price-quantity pairs $\{p_i^*(Z), q_i^*(p_i^*(Z))\}$ and instruments (supply shifters), Z, to approximate the true demand elasticity. The frequently used approach in such cases is to use 2SLS (two-stage least squares) and first approximate the

[5] An alternative interpretation to having n markets that trade homogeneous goods is that the same market clears n times.

price given the supply shifters:[6]

$$p_i^* = Z\pi + \eta, \tag{2.2}$$

where π are the first-stage coefficients, η_i is the first-stage error term, and Z are supply shifters that cause exogenous shocks to the equilibrium price, p_i^*. As common, instruments are assumed to be relevant and independent, i.e., each supply shifter $z_j \in Z$ should be uncorrelated with the second-stage error term, $\text{cov}(z_j, \xi_i) = 0$, but correlated with the equilibrium price, $\text{cov}(p_i^*, z_j) > 0$. Note that the fitted values p_i^* depend on the choice of instruments Z and the choice of weights π. Below, we employ machine learning methods to estimate (2.2) and study how they affect the fitted values p_i^* and, in turn, the elasticity as approximated by the second stage equation:

$$q_i^* = \alpha + p_i^* \epsilon^* + X\beta + \xi. \tag{2.3}$$

In this equation, β captures coefficients of all variables in X, this is $\beta = (\beta_1, \ldots, \beta_m)$ and $X = [x_1, \ldots, x_m]$ captures observable heterogeneity other than price that determines demand in market i. The 2SLS estimate of the true elasticity given the equilibrium price-quantity pairs is represented by ϵ^*.

Note that the 2SLS approach addresses the endogeneity concerns that arise from directly estimating quantities q^* by regressing these on prices p^* using OLS (ordinary least squares). As well known, the equilibrium price is endogenous because it depends on shifts in demand, leading to biased ordinary least squares (OLS) estimates.

Below, we estimate demand elasticity based on observed demand curves as in Equation (2.1), which provides our benchmark estimate for the true underlying elasticity. We also estimate elasticity using equilibrium price-quantity pairs following the canonical IV approach as in Equations (2.2) and (2.3). When using instrumental variables, we employ the standard two-stage least squares approach and machine learning approaches that first select the most relevant instruments from the set of supply shifters Z. The goal is to arrive at good elasticity estimates when demand curves are not observed and only price-quantity pairs are available. In particular, we are interested in the usefulness of machine learning tools for selecting instruments and arriving at better demand estimates.

[6]Note that we use p^* and q^* to denote both, the theoretical equilibrium price-quantity tuples and the underlying data.

2.3 Data

To explore the usefulness of machine learning for estimating demand, our empirical setup exploits data from the European Power Exchange EPEX. In Europe, EPEX acts as the largest power exchange and clears, amongst others, day-ahead wholesale markets for electricity in all major European countries. In this study, we employ data from the German day-ahead electricity wholesale market.

The EPEX organizes market clearing as a closed bid double auction, that determines a uniform clearing price for every delivery hour of the consecutive day. Prior to market clearing, producers and retailers submit stepwise supply and demand schedules, respectively, for each hour of the next day. EPEX aggregates and matches supply and demand, and establishes 24 hourly day-ahead clearing prices. Finally, note that there exist lower and upper bounds for clearing prices. Allowed supply and demand schedules can range between -500 EUR/MWh and 3000 EUR/MWh.

We make use of rich power market data for several reasons. First, electricity is a homogeneous good and allows us to abstract from competing differentiated products. Second, our setup offers a well-defined test environment for demand estimation, because aggregate demand functions, and hence true underlying demand elasticities, are directly observable in the data. We therefore can benchmark estimated elasticities against observed demand curves submitted to the wholesale double auction. Third, we in addition observe equilibrium price-quantity pairs at hourly granularity, which allows us to compute two-stage IV estimates for demand elasticity in a high-frequency setting. Last, there exist several plausible instruments as supply shifters.

The instruments that we employ in our study are weather-dependent day-ahead solar and wind output forecasts which are available in hourly resolution. Because solar and wind output have marginal costs close to zero, both are exogenous supply shifters that shift the entire market supply. We also make use of day-ahead input prices for gas, the front month futures prices for coal, and prices for CO_2 allowances. All latter are inputs to producing electricity and can be considered as exogenous. In total, we hence make use of $k = 5$ first-stage instruments.

As we use solar output forecasts, we restrict our sample to 12 daytime hours between 8 am and 8 pm. Our observation period spans 11 years, from 2010 to 2020. In total, our data consists of 34,440 hourly observations.[7] For each hour, our data includes aggregated demand curves, market clearing price-quantity pairs, and observations on our instrumental variables.

Specifically, the instruments solar output and wind output are observable at the hourly level, while fuel prices and CO_2 prices are at the daily level. Table 2.1 presents selected summary statistics.

Table 2.1: Summary statistics

	Mean	St. Dev.	Min	Max	Obs.
Price (EUR/MWh)	46.07	16.00	-83.94	210.00	34,440
Quantity (GW)	29.94	5.05	16.57	51.47	34,440
Wind output (GW)	8.64	8.04	0.22	47.09	34,440
Solar output (GW)	7.47	7.10	0	32.48	30,960
Gas price (EUR/MWh)	19.16	5.56	3.63	29.35	2,870
Coal price (EUR/metric ton)	9.17	2.04	5.08	14.41	2,870
CO_2 price (EUR/metric ton)	11.85	7.51	2.75	33.44	2,870

Notes: Prices and quantities are for day-ahead electricity at the German market, traded at EPEX. Gas prices are day-ahead prices for the TTF hub. Coal prices are for imported coal to Europe (API2-CIF), CO_2 prices are from ICE. Prices, quantities, wind and solar output each contain one missing observation due to daylight saving time adjustments. The amount of peak hours in each year of our dataset is given by 3,132 for the years 2010, 2012-2016, 2018-2019. For the year 2011 and 2017, we have 3,120 observations. For the year 2020, we have 3,144 observations. Solar output contains 3,480 missing observations. We observe gas, coal and CO_2 prices only in daily resolution.

2.4 Empirical Strategy

In this section, we introduce our empirical strategy to estimating "true" demand elasticities from observed demand curves. Subsequently, we illustrate how we infer the same elasticity from market clearing price-quantity pairs and applying IV-based estimators, with and without machine learning in the first stage. We also present how we evaluate the different estimators.

2.4.1 Estimating True Demand from Bid Curves

We start from the Equation (2.1) and assume that demand is iso-elastic. Then, the demand function can be written as:

$$Q_i = \alpha p^{\epsilon_i}. \tag{2.4}$$

[7] More specifically, the data in total comprise 12 hours between 8 am and 8 pm * 52 weeks * 5 working days * 11 years = 34,320 observations, plus additional hours from leap years and years with 53 calendar weeks.

Note that in our data, market i refers to a specific market clearing in hour i. The demand elasticity in hour i is then directly estimable by using data points along the submitted demand curves, using a log-log-specification:

$$\log(Q_i) = \log(\alpha) + \epsilon_i \log(p) + \xi, \tag{2.5}$$

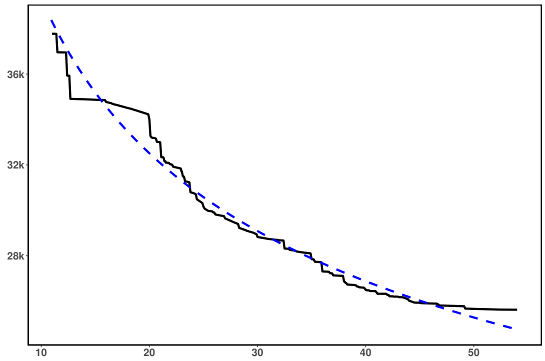
where ϵ_i is the estimate of the elasticity in market i as in Equation (2.1). In essence, the above fits an iso-elastic curve around the observed demand function. We apply Equation (2.5) to all hours in our sample and compute the underlying demand elasticity for every hour i by averaging over all markets' elasticities belonging to that hour. As this estimate is obtained directly from using demand curve data, we also refer to this estimate as "true" elasticity and denote it by ϵ_i^{true} henceforth.

Note again that our observed demand curves are stepwise functions. The steps of the submitted curve are well-approximated by an iso-elastic function with an average approximation error of only 2.24% for our entire sample. Figure 2.1 depicts four representative hours from our sample with their corresponding "true" elasticity estimate ϵ_i^{true} as estimated by Equation (2.5). As can be seen, the assumption of iso-elastic demand fits the data well. As indicated in Figure 2.1, we estimate Equation (2.5) only using relevant parts of the demand curve, i.e., we fit the log-log specification up to a specified highest observed equilibrium price to avoid fitting demand curves significantly out of sample.[8] Appendix A.1.1 presents this approach in more detail. In Table 2.2, we illustrate the summary statistics of the resulting estimates. We plot selected statistics of the true elasticity by year. As can be seen, the "true" elasticity varies relatively little throughout the year with a typical standard deviation of below 5%.

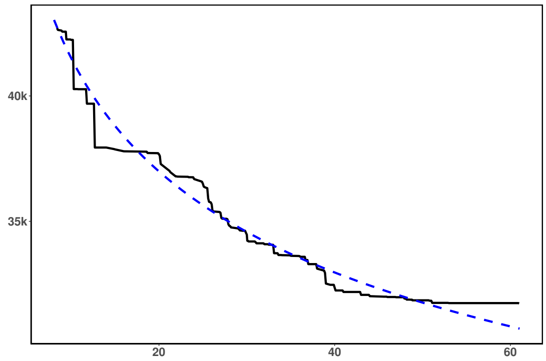
Lastly, to make our estimates comparable with those derived from equilibrium price-quantity pairs, we average the true elasticities over the same timespan as the sample used for our IV estimations. For this sake, we partition our 11-year sample in different subsamples ω, for each of which we then compute the average elasticity by averaging over estimates from (2.5). Formally,

$$\epsilon_\omega^{true} = \frac{1}{n_\omega} \sum_{i \in \omega} \epsilon_i^{true}, \tag{2.6}$$

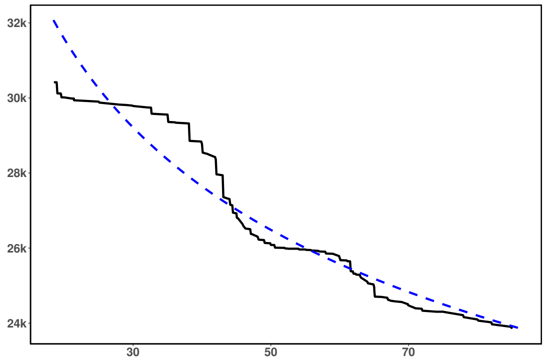
[8] Recall that demand can be measured for all prices between -500 and 3000 EUR/MWh but that clearing prices as reported in Table 2.1 range between -83.94 and 210 EUR/MWh. Importantly we include negative prices when estimating Equation (2.5) by transforming the prices with the function $\text{sign}(p)\log(1 + |p|)$. This transformation preserves the sign of the prices as opposed to leaving negative prices out of the sample and significantly improves the fit as compared to the untransformed observed demand curve. Our results are robust with respect to this choice of the price transformation.



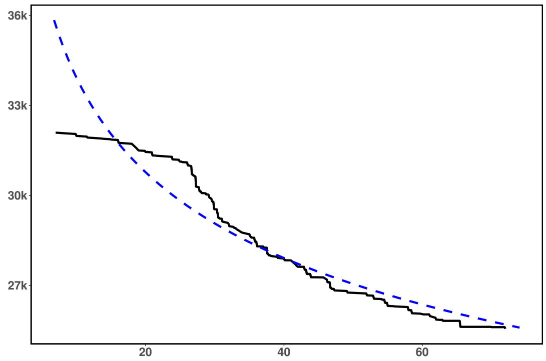
(a) Hour 9, September 1^{st}, 2016



(b) Hour 15, March 16^{th}, 2017



(c) Hour 8, August 10^{th}, 2018



(d) Hour 10, July 11^{th}, 2019

Figure 2.1: Submitted demand curves (quantity in thousands of MW over price in EUR/MWh) and elasticity ϵ_i^{true}. Part (a) of this figure shows the observed stepwise demand curve for the hour 9-10am on September 1^{st}, 2016. The dashed line shows the fitted iso-elastic demand curve with elasticity ϵ_i^{true} that we obtain when estimating the log-log specification in Equation (2.5). Parts (b), (c), and (d) show demand curves and corresponding log-log fits for additional three representative hours.

where ϵ_ω^{true} denotes our estimate for the underlying average elasticities that we obtain from the set of observed demand curves for the sub-sample ω. For instance, when looking at weekly average elasticity, the sample ω consists of all hours belonging to a given calendar week.

This strategy allows us to compare different IV estimators against the true elasticity for different sample lengths and sample compositions. For example, for week-level aggregation, we pool all calendar weeks into one sample and perform the above procedure. This is, we fit iso-elastic demand curves for each hour of every calendar week 20 in the sample and obtain ϵ_ω^{true} for this particular calendar week. Subsequently, when using all hourly price-quantity pairs and instruments from calendar week 20 (again from 2010 to 2020), we then estimate Equations (2.8) and (2.7) to obtain the IV estimates. We then evaluate our IV estimates against ϵ_ω^{true} for week 20. Our main analysis is performed at the weekly level. Yet, for additional tests, we also pool all markets for every peak hour over our entire sample into one sub-sample per peak hour and

Table 2.2: Summary statistics: true elasticities of peak hours (hours 8 to 20)

Year	Mean	St. Dev.	Min	Max	Obs.
2010	-0.299	0.111	-0.766	-0.060	3120
2011	-0.279	0.093	-0.801	0.064	3101
2012	-0.197	0.072	-0.490	-0.038	3125
2013	-0.123	0.050	-0.322	-0.023	3125
2014	-0.192	0.069	-0.438	-0.043	3125
2015	-0.224	0.073	-0.484	-0.042	3125
2016	-0.229	0.108	-0.600	-0.015	3125
2017	-0.174	0.095	-0.574	-0.015	3125
2018	-0.204	0.144	-0.823	-0.010	3125
2019	-0.139	0.084	-0.476	-0.080	3125
2020	-0.101	0.067	-0.352	-0.060	3125

Notes: Only hours with complete information on demand curves are included: the differences in observations compared to the Table 2.1 are due to missing demand curves in our sample. Standard deviations present sample estimates. Only peak hours are considered, i.e. Mon-Fri, hour 8 to 20.

then approximate the average elasticity of that peak hour only, again using IV estimates. We then compare hourly IV estimates to the corresponding true demand elasticities.

2.4.2 Estimating Demand Elasticity from Price-quantity Pairs

Next, we describe our approach to estimate elasticity from market-clearing price-quantity pairs. We follow the approach from Section 2.2 and apply log-log specifications throughout. In the first stage of the estimation, we approximate equilibrium prices p_i^* via supply shifters Z by fitting:

$$\log(p_i^*) = \log(\delta) + \sum_{j=1}^{k} \pi_j \log(z_j) + \sum_{l=1}^{m} \pi_l x_l + \eta, \tag{2.7}$$

where $z_j \in Z$ are supply shifters as introduced in Section 2.2. Note that exogenous controls $x_l \in X$ also appear in this equation, as they ensure consistency of the 2SLS estimator in the second stage. X consists of m controls related to time, i.e., depending on the specification, we employ hourly, daily, weekly and monthly dummies. We also use the day of the week as a dummy in hourly models.

Next, we extract fitted values from this estimation and denote these by $\tilde{p}_i^*$. These fitted values are then substituted into the second stage:

$$\log(q_i^*) = \log(\alpha) + \epsilon_i^* \log(\tilde{p}_i^*) + X\beta + \xi, \tag{2.8}$$

where q_i^* are equilibrium quantities and $\tilde{p}_i^*$ are fitted equilibrium prices in market i. ϵ_i^* is our IV-estimate for the average elasticity across all markets, ϵ.[9] In our data, q_i^*, p_i^* are the prices and quantities which clear the double auction.

The resulting elasticity estimate ϵ^* can then be compared to the true elasticity stemming from (2.6) for evaluating the quality of inference. Of course, the effectiveness of this method heavily relies on the instruments being strong and independent of the error term. When this condition is not fulfilled or only fulfilled partly, the resulting estimate from the two-stage procedure can become even worse than a regular OLS estimate.

2.4.3 Regularized First-stage Estimation

To select the best instruments in the first stage, we therefore apply different strategies to obtain fitted values $\log(\tilde{p}_i^*)$. Below, we briefly introduce methods from the machine learning literature that we employ. Specifically, we introduce regularization into the first-stage Equation (2.7) and study how regularization affects the second-stage estimate.

First, we denote the regularized first-stage estimator with L1 penalty as π_L. To compute the latter, we solve:

$$\pi_L \in \underset{\pi \in \mathbb{R}^{k+m}}{\arg\min}\, \mathbb{E}[(\log(p_i^*) - (\log(Z), X)\pi)^2] + \frac{\lambda}{n} \|\pi\|_1 \,, \tag{2.9}$$

where λ is chosen such that the 10-fold cross-validated MSE is minimal. Note that π_L has $k+m$ components, as we regress on both, k instruments and m exogenous variables and controls X in the first stage.

Following the same notation, the L2-penalized first-stage estimator (RIDGE) is defined as:

$$\pi_R \in \underset{\pi \in \mathbb{R}^{k+m}}{\arg\min}\, \mathbb{E}[(\log(p_i^*) - (\log(Z), X)\pi)^2] + \frac{\lambda}{n} \|\pi\|_2^2 \,, \tag{2.10}$$

[9] In abuse of notation, we use ξ_i as error term for estimating demand from bid curves and price-quantity pairs.

and again, the regularization intensity λ is chosen to minimize MSE using 10-fold cross-validation.

Last, we construct the Post-LASSO estimator π_{PL} by first performing a LASSO-penalty regularization in the first stage and then taking all non-zero coefficients to estimate the second stage equation using OLS, formally:

$$\pi_{PL} \in \underset{\pi \in \mathbb{R}^{k+m}}{\arg\min}\, \mathbb{E}[(\log(p_i^*) - (\log(Z), X)\pi)^2] : \pi_j = 0 \ \forall \ j : \pi_j^L = 0. \tag{2.11}$$

The idea for this estimator is similar to Belloni et al. (2012), who also apply LASSO as a pre-selection procedure.

The regularization parameter λ, if positive, introduces first-stage bias as compared to the standard 2SLS case with $\lambda = 0$. As λ increases, so do the costs of large instrument coefficients when minimizing the sum of squared errors.

The above setup allows for obtaining several estimators of ϵ^* for the second-stage model in Equation (2.8). First, using fitted values for p_i^* with $\lambda = 0$ yields the standard two-stage least squares estimator ϵ^*_{2ls}. In addition, we obtain three regularized IV estimators when using fitted values from (2.9), (2.10) and (2.11), i.e., using LASSO, RIDGE or Post-LASSO penalties. We denote the two-stage estimator based on RIDGE first-stage as ϵ^*_R, the second-stage estimator based on LASSO regression as ϵ^*_L and the Post-LASSO second-stage as ϵ^*_{PL}.

Apart from investigating how regularized first-stage estimators choose instruments and their weights, we will also look into the best data-driven choice for the regularization intensity λ in terms of bias and MSE in the second stage.

2.4.4 Evaluation Criteria

Typically, researchers who estimate demand elasticity from price-quantity pairs (or apply IV estimation to other problems) are interested in approximating the true elasticity (effect of interest) as closely as possible, while keeping the variance of the estimates low. For prediction purposes, the mean squared error (MSE) becomes relevant, too. We therefore compare and report the performance of estimators based on the bias, variance, and MSE. For each sample ω, we define and compute the sample analogs of bias, MSE, and standard deviation as

$$Bias(\epsilon^*_\omega) = \epsilon^{true}_\omega - \epsilon^*_\omega, \tag{2.12}$$

$$MSE(\epsilon_\omega^*) = \sum[(\epsilon_\omega - \epsilon_\omega^{true})^2], \tag{2.13}$$

$$\sigma(\epsilon_\omega^*) = \sqrt{\frac{\sum(\epsilon_\omega - \epsilon_\omega^{true})^2}{N}}, \tag{2.14}$$

where N denotes the total number of market clearing hours i in sample ω.

Last, we also investigate how the regularized methods choose instruments' weights and which coefficients are set to 0 by the LASSO estimator. For each pool of markets, we consider which instruments are being left in the sample and provide economic reasoning for the type of selection behavior.

2.5 Results

We first present results for estimating weekly average elasticities. This is, we pool hourly markets at the weekly level over the entire timespan of the dataset and estimate ϵ_ω^{true} and $\{\epsilon^*_{2ls,R,L,PL}\}$ for the corresponding weekly samples. As we have 11 years of data, the estimates are hence based on 11 full-week observations for each of the 52 calendar weeks. Figure 2.2 displays the bias over all calendar weeks for all considered estimators. As shown, the bias reduces considerably and estimates vary much less for the regularized first-stage models.

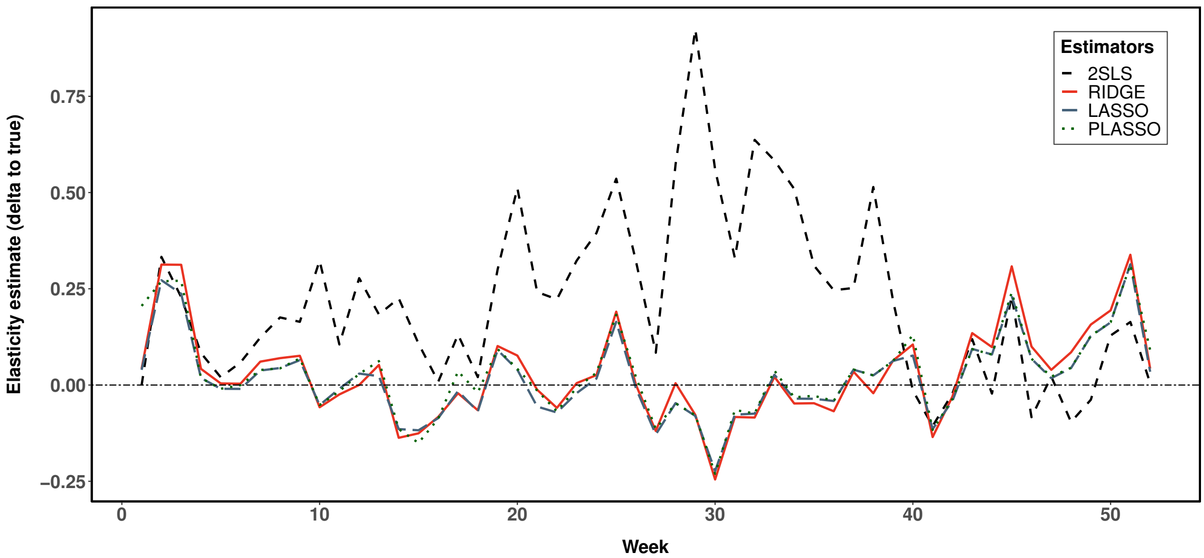


Figure 2.2: Bias of two-stage estimators (2SLS, RIDGE, LASSO, Post-LASSO). This figure shows the bias of each estimator for each of the 52 calendar weeks. The bias is shown relative to the true demand elasticity for the respective week. Note that we plot the bias as stated in Equation (2.12), but swap the estimated elasticity and the true elasticity. This allows for a more natural interpretation of the plot, as positive values represent an overestimation while negative ones represent an underestimation.

We observe that the relative over-performance of regularized estimators is especially pronounced in weeks with very high variance in instruments, especially for solar output during summer weeks. In these cases, all three of the regularized estimators have a lower bias and return more stable results.

Table 2.3 reports the mean bias, mean squared error, and variance of the estimators. As shown, the estimation quality improves considerably for our regularized IV estimators, and two-stage estimates based on LASSO perform best.

Table 2.3: Bias, mean squarer error, and variance of two-stage estimators over 52 weeks

Estimator	Bias	MSE	Variance	Observations
ϵ^*_{2ls}	0.221	0.428	0.218	31,100
ϵ^*_{R}	0.029	0.088	0.104	31,090
ϵ^*_{L}	0.018	0.064	0.109	31,090
ϵ^*_{PL}	0.030	0.087	0.108	31,090

Notes: Bias, Variance and MSE are computed as in Equations (2.12), (2.13), and (2.14). The differences in the numbers of observations compared to Table 2.1 are due to daylight saving time shift and erroneous data. In particular, we require complete observations on prices, quantities and *all* instruments to compute these estimators. Therefore, for a given 15 minutes interval, one missing observation leads to exclusion from our sample.

Regularization is commonly known to outperform unregularized estimators, particularly in cases of a small number of observations compared to the number of explaining variables, collinearity between explanatory variables, and a high amount of outliers. While price-quantity datasets usually exhibit enough observations, collinearity and outliers are a concern. Indeed, we find variance inflation factors (VIF) of up to 8 in our data on instruments, which is a clear sign of collinearity. Also, fundamentally speaking, fuel prices are typically highly correlated. When investigating outliers, we find that, on average, around 3% of observations are outliers in terms of Cook's distance. We believe this contributes to the performance of the regularized estimators, as shown in Table 2.3.

We check for the robustness of our results by pooling the data by peak hours instead of weeks. This analysis also shows how the number of observations can influence inference precision. We therefore compare ϵ^{true}_{ω} and $\{\epsilon^*_{2ls,R,L,PL}\}$ where a each sample ω now contains all peak hours of hour i with $i \in \{9, 20\}$ over all years of our dataset. Figure 2.3 displays the difference between the true elasticity and our estimates based on price-quantity pairs. Again we find that the classical IV approach largely underestimates the hourly demand elasticity, while all of the regularized estimators perform significantly better. On average, the bias decreases by around 50%. The average bias of the 2SLS estimator is around -0.095, while the average bias for the regularized estimators is between -0.041 and -0.043.

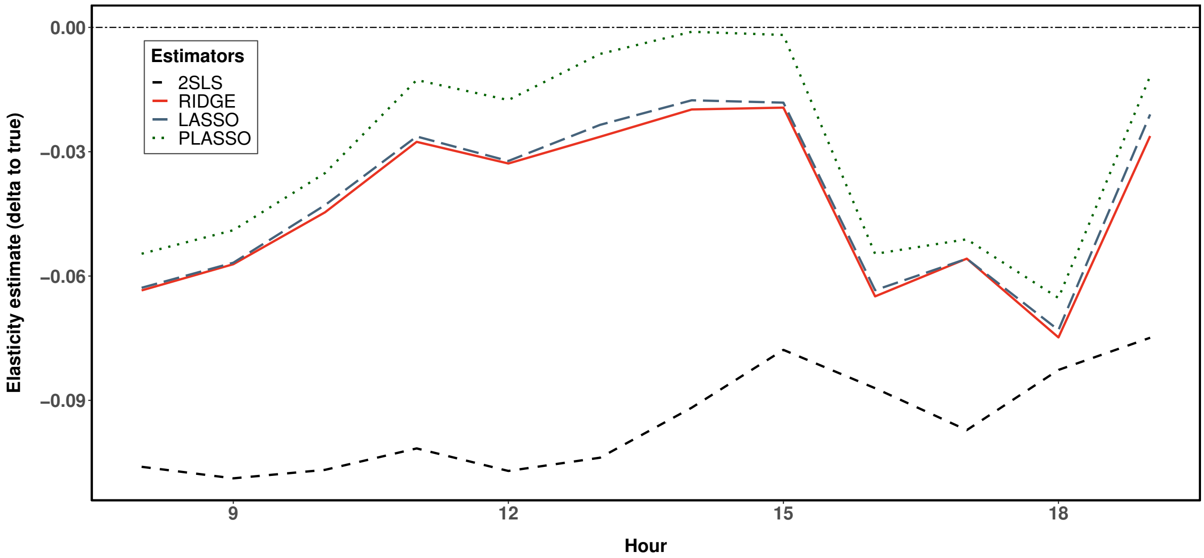


Figure 2.3: Bias of two-stage estimators (2SLS, RIDGE, LASSO, Post-LASSO). This figure shows the bias of each estimator for each of the 12 peak hours. The bias is shown relative to the true demand elasticity for the respective hour.

To investigate how the observed improvements are related to instrument selection, we look into which instruments are being selected over time by the regularized estimators. In case of LASSO and Post-LASSO, this can be done by screening for instruments with coefficients of zero. We find that in our weekly models, the instruments coal, gas and CO_2 prices are set to zero in 6%, 20% and 14% of cases, respectively, while wind and solar remain relevant almost always. In addition, the relevance of the instruments changes over time. As we start in 2010 and analyze the timespan until 2020 where wind and solar heavily entered the market, we conjecture that the explanatory power of the forecasted renewable energy in-feed should increase over time. Table 2.4 reports how often each instrument is selected in the first stage over the years. We find that while wind feed-in is kept in all models, the role of solar feed-in in explaining first-stage price variance is increasing over time. Coal and gas prices remain at relatively low levels in terms of explaining price. Notably, the CO_2 price gains more explanatory power over time.

Last, we present the sensitivity of our results with respect to the level of regularization. The choice of λ in regularized models is subject to ongoing research. We find that the performance of regularized estimators can drop significantly when the level of applied shrinkage becomes too high. Without the possibility to benchmark against true elasticities in the second-stage, the task

Table 2.4: Instrument relevance over time: percentage of models which use the corresponding instrument in the first stage after LASSO-regularization

Year	Solar	Wind	Coal pr.	Gas pr.	CO_2 pr.
2010	0.83	1.00	0.58	0.58	0.66
2011	0.83	1.00	0.58	0.83	0.75
2012	0.92	1.00	0.83	0.83	0.83
2013	0.92	1.00	0.75	0.83	0.83
2014	0.92	1.00	0.50	0.91	0.58
2015	1.00	1.00	0.50	0.58	0.91
2016	1.00	1.00	0.66	0.41	0.58
2017	1.00	1.00	0.50	0.50	0.91
2018	0.92	1.00	0.50	0.66	0.91
2019	1.00	1.00	0.50	0.58	0.91
2020	1.00	1.00	0.50	0.58	0.91

Notes: For each year, we consider how many estimation models per month are using the corresponding instrument variable after regularizing the first-stage and express this amount as percentage.

of choosing the shrinkage level correctly becomes non-trivial. We find that the standard 10-fold cross-validation with respect to the first-stage MSE as implemented in most statistical packages yields a too high regularization level in about 20% of the cases. This means that in these cases, reducing the level of shrinkage yielded a better second-stage estimate. Furthermore, we find that if using cross-validation, the fold size should be adjusted to the number of observations and is highly sensitive with respect to it. For example, when fitting weekly models with about 60 observations, reducing the fold number to 3 yielded the best results.

2.6 Conclusion

The increasing availability of large datasets has led to the growing prominence of statistical learning methods in empirical economics. While these methods have originally been developed to improve out-of-sample prediction and MSE, they are increasingly adopted and used to improve inference.

In this article, we make use of a unique empirical setting to put these methods to a test. We contribute to the empirical economics literature by showing how regularization in the first-stage of the canonical demand estimation problem can help to improve the second-stage inference.

We first define the true demand elasticity as the one derived from the submitted bid curves. This elasticity estimate directly reflects the price response of the demand side and we therefore use it as a benchmark. We then define estimators which can approximate this elasticity from information on equilibrium prices-quantity pairs and supply shifters as instruments. We show how regularization in the first-stage is used to address the instrument selection and the weak instrument problems.

Using this empirical strategy, we derive a set of results on the properties of RIDGE, LASSO, and Post-LASSO estimators. We find that bias, variance, and MSE of the second-stage coefficients can be significantly reduced when employing regularization in the first-stage. The bias reduction property is more prominent the higher the number of observations and instruments. We also find that the variance of the estimates is reduced by more than 50%.

We find that the observed effects are robust with regard to the aggregation level and amount of observations. We propose several explanations for the better performance of regularized estimators. We find that the number of outliers in the observations, the collinearity of instruments, and the changing relevance of instruments over time are all factors explaining the performance of regularization.

In addition, we investigate how the level of shrinkage affects the results. We find that the estimators are quite sensitive with regard to the selection of the shrinkage parameter λ. We argue that "over-shrinking" is a concern, as the quality of estimates can drop significantly when it occurs. We derive suggestions on the practical choice of the shrinkage parameter based on our observations.

In sum, our findings suggest that regularization methods offer measurable advantages for IV estimation with large data and many instrument candidates, with LASSO-type regularization performing best on our data.

3 | Optimal Liquidation in Continuous Markets with Workups

Vadim Gorski, Sebastian Schwenen[10]

Traders active in financial markets based on limit order books can choose between submitting limit or market orders to liquidate their position. However, they have an incentive to shade their true bids in fear of price impact. This effect has led to the emergence of dark pools and quantity-based allocation mechanisms such as size discovery sessions. We develop a model of a limit order market, unifying market orders and limit orders with a randomized size discovery mechanism. Based on this model, we study the optimal liquidation problem and resulting allocative efficiency. We find that introducing randomized size discovery sessions increases the reallocation speed, allows market participants to avoid costly holding of unwanted inventory, and thereby increases welfare. We show this effect on historical LOB data from the Nasdaq Nordic power market. We argue that while this mechanism is attractive in theory, practical implementation bears some problems.

[10] Author contributions: This essay is based on a joint paper with Sebastian Schwenen. My contribution was the development of the research idea, large parts of the theoretical model, the empirical work, and major parts of the draft of this paper

3.1 Introduction

Today, continuous trading on lit exchanges[11] using limit order books (LOB) represents the most important allocation mechanism in financial markets. Traders active in LOB markets use a variety of strategies to optimally liquidate their undesired inventory. Especially for traders with large positions, the liquidation strategy plays a key role, as they face significant adverse price effects when liquidating.

Strategic behavior in LOB markets is the reason why liquidating a large position becomes challenging. Traders placing large orders are facing an uncertain price impact and are susceptible to front-running. Front-runners use the information in the LOB to position themselves against traders with large positions (Pancs, 2014). This and other kinds of strategic behavior also impede an efficient price discovery (Vayanos, 1999). The consequences of this problem are two-fold.

First, traders typically split up their true position into smaller pieces or employ iceberg orders, thereby shading their true bids. While splitting-up orders can be an effective technique, it slows down the reallocation speed, i.e., it takes market participants longer to adjust their portfolio and arrive at the allocation of financial assets they require. Splitting up the entire position into smaller ones also has a substantial impact on transaction costs, adding complexity to the allocation decision. Additional adverse effects of splitting are costly delays in the reallocation of financial assets (Rostek and Weretka, 2015).

Second, exchanges with non-public order books, so-called dark pools have emerged as a response. They offer traders the possibility to exchange large positions without price impact, because trades can be submitted privately. This has made dark pools ubiquitous in many financial markets. Dark pools offer a wide variety of trading mechanisms, with size discovery sessions being one of them. Despite their popularity (Fleming et al., 2018), both the empirical and the theoretical literature have no clear answer on the role of dark pools and, in particular, size discovery mechanisms for the efficient reallocation of financial assets and welfare effects. We refer to size discovery sessions as workups, as size discovery is a particular workup mechanism.[12] While there are seminal results on the role of workups (Pancs, 2014, Duffie and Zhu, 2017) and on the optimal liquidation with market and/or limit orders (Almgren and Chriss, 2001, Foucault et al., 2005, Cartea and Jaimungal, 2015, Roşu, 2019), the literature so far is largely silent on

[11] Lit exchanges are venues with publicly visible bids and offers.
[12] Generally, a sequential process of negotiating on a quantity is called "workup".

the combination of these mechanisms into one liquidation strategy, albeit their joint use in many markets.

In this paper, we address the role of size discovery in LOB markets from two perspectives. From the perspective of a trader, we solve the optimal liquidation problem when size discovery sessions are offered in addition to regular LOB trading. From the perspective of an exchange, we derive the incentives for traders active in such markets, thereby shedding light on the interplay between a lit market LOB and a size discovery session offered by a dark pool. This has implications for the strategy of exchanges looking to attract traders.

Our model for optimal liquidation builds on the classic optimal liquidation model by Almgren and Chriss (2001) in continuous time. In the environment that we study, traders can decide between placing a market order, a limit order with a certain mark-up compared to the market price, or placing a part of the volume in a workup session. All three decisions are made jointly, i.e., traders optimize their strategy over all three options at all times.

The trade-off traders face is the penalty to hold unwanted inventory versus the penalty of a market impact when they execute too large positions too quickly. The market impact effect can be reduced by either using limit orders or by participating in workups. Both options, however, do not guarantee execution. Instead, we introduce a probability of a position being executed given a limit order with a certain mark-up and stochastic liquidity available in a workup session. The trade-off between the price impact of a market order and the uncertainty of execution is another key trade-off in our model. The price impact in our model is linear, which is a classic choice in the literature (Huberman and Stanzl, 2004).

In addition, we address the problem of strategic behavior in markets with workups by randomizing the arrival of liquidity in a workup session. This mechanism addresses the problem of traders withholding liquidity and strategically waiting for workups, thereby harming price discovery in lit markets.

The first key result of our paper is the optimal liquidation choice of a trader. The parameters in our model can all be estimated from historical LOB data, thereby making it applicable in practice. The basic version of our model allows us to find explicit, closed-form solutions. Estimated parameters can be substituted into these solutions to find the optimal strategy. An extension of our model incorporates decisions by other traders via order flow (Bechler and

Ludkovski, 2015), but does not offer a well-tractable analytical solution. Yet, we show how the solution can be approximated numerically.

Our second key result is that the strategies for trading in lit markets and dark pools cannot be optimized separately. We find that the effect of participating in workups has an influence on the optimal placement of market and limit orders on a lit exchange. Hence the order strategies are endogenous to participating in workups.

We also apply our model to LOB market data. In particular, our application relies on data from Nasdaq Nordic power market. First, we find that following the optimal lit market strategy results in substantially lower price variance and less adverse price movements as implementing benchmark strategies such as time- or volume-weighted order placement. Second, we show in a simulation-based on real order-book data that a workup session introduced at random times increases the trading volume and thereby the reallocation speed with only little adverse effect on the lit market liquidity.

Overall, this paper provides a theoretical model to analyze trading in lit markets and dark pools, a set of empirical findings for power markets, and a set of empirical strategies for analyzing LOB data. As such, we contribute to a growing literature on optimal liquidation, dark pools, workups, and the empirics of LOBs.

As shown in Fleming et al. (2018), around 50% of the total volume traded in US-American treasuries is exchanged in workup sessions. For corporate bonds and credit default swaps (CDS) products, around 70% are exchanged over so-called matching sessions with a fixed price set by an operator. In addition, around 15% of US equities are traded over dark pools. Companies like BrokerTec are even offering randomly held workup sessions.

Starting with Pancs (2014), the literature has studied the effect of workups on trading strategies theoretically. In particular, he focuses on expandable orders as one form of quantity negotiations. He compares expandable orders to iceberg orders and identifies front-running and lack of trust in an exchange as common problems that incentivize traders to participate in workups. He proposes a "button mechanism" which allows traders to gradually submit their quantities to avoid information leakage. This is structurally very similar to the size discovery mechanism proposed in our work. However, our theoretical model is different, as we consider this problem from the optimal liquidation perspective.

Duffie and Zhu (2017) introduce a theoretical model of workups in the form of size discovery sessions. Their model shows that under some conditions, a one-off size discovery session at the beginning of the trading day is more efficient than a regular LOB market. This idea is similar to holding an opening and closing auction-based mechanism in addition to regular LOB trading (Friederich and Payne, 2007, Huang and Tsai, 2008). Their results, however, only hold under common value assumption and uniform holding costs. In addition, they do not model the relationship between participation in a workup session and trading in the lit market. The drawback of workup sessions they identify is that traders anticipating the session have ample incentives to wait for it, thereby withholding liquidity.

To avoid strategic waiting for the workup, the arrival of the session can be randomized in time, an idea related to Antill and Duffie (2019) who introduce a size discovery session which arrives at random times. In this model, traders submit an optimal demand function and an excess inventory for the size discovery session. Both are then processed by the exchange. In contrast, our model allows traders to decide between different modes of trading to reduce their unwanted inventory. In addition, Antill and Duffie (2019) argue that size discovery harms overall welfare by reducing liquidity in lit markets. While this result holds true in a batched auctions setting, as, e.g., proposed by Budish et al. (2015), it is not clear how this effect influences a continuous LOB market that we model in our paper. Du and Zhu (2017) introduce a related model, where traders also submit demand schedules. They use the model to find an optimal frequency of a sequence of double auctions so as to optimize the reallocation speed. While the speed of reallocation is optimized in our model as well, Du and Zhu (2017) do not incorporate size discovery in their model.[13]

Apart from modeling workups, the integration of limit orders into the optimal liquidation strategy complicates the resulting analysis considerably. Parlour and Seppi (2008) provide an overview of the main obstacles of modeling limit orders, such as deciding on price and quantity simultaneously, the interplay of limit and market orders, and the stochasticity of order fills. We address this problem in our model by introducing mark-ups above the mid-price as the decision variable and modeling the filled limit order quantity via a probability distribution representing current market liquidity.

[13]Speed of reallocation plays a role in the empirical literature, too. Riordan and Storkenmaier (2012) find that exchanges offering higher speeds of trading and lower latencies attract liquidity. This supports the idea of reallocation speed being an important metric for traders.

We also contribute to the literature on optimal liquidation modeling by integrating limit orders, market orders, and size discovery sessions into one model. Foucault et al. (2005) model the optimal choice of orders assuming symmetric information and sequentially arriving traders. They do not account for strategic choice between market and limit orders but introduce different levels of waiting costs as one of the drivers of the liquidation decision, a property also used in our model. Roşu (2009) propose a similar model but allow sequentially arriving traders to modify already submitted limit orders. In a more recent work, Roşu (2020) studies a setting with both, market and limit orders in a continuous dynamic model. His focus, however, is on information asymmetry among traders, while we focus on the optimal liquidation, assuming traders follow the same liquidation strategy. Alfonsi et al. (2010) solve an optimal execution problem for a LOB market with an arbitrary shape of the order book. Similarly, Obizhaeva and Wang (2013) and Siu et al. (2019) introduce a detailed, discrete-time model of a limit order book and find that an optimal limit-order strategy mainly depends on the resilience of the order book. The shape of the order book does not play a role in our theoretical model, as we directly impose a distribution of liquidity. Cartea and Jaimungal (2015) introduce a continuous-time model of the optimal liquidation problem with limit and market orders. Their model, however, does not account for dark pools.[14]

In addition to the optimal liquidation literature, other theoretical models consider optimal order routing. This literature relates to our model, as size discovery sessions are typically a dark pool mechanism, with dark pools often operating independently of the lit exchange. Cont and Kukanov (2017) formulate a model with multiple trading venues and solve an optimal order routing problem. Traders face an order placement problem after they have decided on an optimal liquidation schedule, that we derive in our model. Therefore, order placement can be seen as an implementation of the liquidation schedule. Similarly, Kratz and Schöneborn (2018) deal with dark pools as a separate venue and optimize the order routing decision. While they incorporate adverse selection in their model, they do not account for market orders.

Apart from order routing, optimal order placement is another related problem. As Guéant et al. (2012) argue, finding an optimal trading strategy and actually implementing it on the market

[14]Two notable parallel modeling approaches of the optimal liquidation problem are based on game theory and zero-intelligence models. Sannikov et al. (2016) present an alternative, game-theoretic modeling approach. They model the interactions between traders with private information and asymmetric players. An interesting parallel take on this problem is found in the zero intelligence literature assuming random behavior by traders. Some properties such as increased market order placement when spreads are narrow and increased limit order placement when spreads are wide can be derived under no strategic assumptions on traders' behavior (Farmer et al., 2005, Foucault, 2010).

are two separate problems. This decoupling of optimal liquidation and the implementation of the trading strategy requires traders to employ various techniques in order to make the two strategies reflect each other closely. Most notably, this is why iceberg orders are widely used in continuous markets. There is abundant evidence that these decrease allocative efficiency (Frey and Sandås, 2017). We provide empirical evidence that by augmenting the trading mechanism with a workup session, regulators and exchanges can increase reallocation speed and thereby allocative efficiency.

Our result that traders have an incentive to split up liquidity between lit markets and dark pools is largely supported by the empirical literature. Extant empirical works such as Degryse et al. (2015) find that dark pools have a negative impact on the market depth, as traders withdraw their volume from continuous exchanges because of their preference for fixed prices. This, in turn, leads to wider bid-ask spreads on these exchanges and overall less liquidity. Similarly, Johann et al. (2019) find that the impact of size discovery is largely negative, decreasing the liquidity on exchange markets. Nimalendran and Ray (2014) study equity markets to show that the price impact worsens when introducing "unlit" markets. Additional evidence from Hatheway et al. (2017) and Farley et al. (2018) points to negative consequences of unlit markets, in particular, order flow segmentation and decreased overall liquidity in each sub-market. Evidence pertaining to non-equity markets is scarce. There is some limited evidence on the efficiency of price discovery mechanisms in fixed income markets, as found in Brandt et al. (2007) and Putniņš (2013). These, however, make no connection between the price discovery and the value added of size discovery in these markets. Apart from efficiency, there are ample studies on the adverse effects of dark pools pertaining to market manipulation. Mittal (2008) finds that information leakage and information asymmetry are two decisive reasons why dark pools might be harming the market. Last, there is also some experimental evidence on the effects of hiding liquidity in limit order markets. Last, Bloomfield et al. (2015) find that while allowing traders to hide their true position via iceberg orders or dark pool orders influences their strategy, the market outcomes resulting from such mechanisms remain largely unchanged, with price discovery and liquidity flow stabilizing around the same levels. Our empirical findings suggest that there is indeed a negative liquidity effect in the LOB market augmented by workups. However, this negative effect is offset by an overall faster reallocation speed.

The remainder of this paper is organized as follows. Section 3.2, presents our market model and the decision problem of the trader. In Section 3.3, we solve the resulting optimal control problem and derive the properties of the optimal solution. We also present numerical simulations and intuition on comparative statics. In section 3.4 we apply our results to a real LOB dataset and test our model in practice. Section 3.5 concludes.

3.2 Model

We model a continuous double auction market with size discovery sessions. We build on the foundation of the classical Almgren-Chriss market model (Almgren and Chriss, 2001) and extend it to include workup sessions held at a price generated by the price discovery process through continuous trading. We allow traders to place market orders, limit orders, and participate in workup sessions. We draw from models by Duffie and Zhu (2017), Antill and Duffie (2019) and consider the fundamental trade-off between a costly delay in liquidation and the price impact of immediate liquidation. The volume to be liquidated on the market is considered unwanted inventory in our model. We introduce random mechanism components for the fulfillment of limit orders and the liquidity process in the workup session. In an extension to our model, we add order flow (Cartea and Jaimungal, 2016) to account for strategies followed by other (competing) traders. We start by laying out the key components of our model.

3.2.1 Market

Assume a market with one security, traded at mid-prices P_t which follow a stochastic process.[15] In the following, we consider the case of arithmetic Brownian motion with drift as our mid-price process. The time is continuous in our model. The beginning of the trading session is denoted by t_0, and the session lasts until T.

Similar to Du and Zhu (2017), we assume that traders aim to liquidate a non-trivially sized portion of their portfolio. The portion is supposed to be non-trivial in the sense that it can cause substantial adverse price effects when liquidated at once. The associated inventory process is denoted by Z_t. The trader aims at reaching his target inventory z^* by the end of the trading session: $Z_T = z^*$. Without loss of generality, we assume $z^* = 0$ and $Z_t > 0$, such that the trader aims to eliminate the entire unwanted position by the end of the trading period and acts as a net seller. We do not explicitly forbid the seller to buy shares in our model, i.e., we do not

[15] We use capital letters for stochastic processes, and we use small letters for a particular value of that process in time.

introduce any constraints to our model here. Our results do not depend on the market side, as both are treated symmetrically.

At each point in time prior to the end of the trading session, a trader holds an unwanted allocation, $Z_t > 0$. Each trader seeks to liquidate such that his utility is maximized. This is equivalent to finding a trading strategy S_t maximizing the utility of the trader. The utility function is typically chosen to reflect both, the current state of the inventory and the disutility of too high inventory at the end of the planning period. The strategy S_t consists of three distinct decisions (controls) a trader can make over time.

The first part is placing a market order at a rate $m \equiv (m_t)_{0 \leq t \leq T}$. Intuitively, this is the number of shares the trader is continuously trying to liquidate via market orders.

The second part is placing a limit order with a mark-up of δ_t above the current mid-price and a size of $l(\delta_t) = l_t$, denoted by $\delta \equiv (\delta_t)_{0 \leq t \leq T}$. We introduce $l(\delta_t)$, because the control for limit orders is stated in terms of *price mark-up* and not in terms of quantity. The function $l(\delta_t)$ maps the chosen mark-up to a quantity posted at that mark-up. In our basic model (Section 3.2), we set $l(\delta_t) = a$, meaning that the trader always aims to liquidate a shares at once. In our model extensions, we set $l(\delta_t) = 1$ for analytical tractability.

Third, traders can choose to participate in a workup session with the posted inventory denoted by $w \equiv (w_t)_{0 \leq t \leq T}$. The total trading activity in terms of shares submitted for transaction at time t can thus be described as:

$$S_t = m_t + l_t + w_t, \tag{3.1}$$

and we introduce $u \equiv \{m, \delta, w\}$ as a set of all controls.[16] Note that the fraction of l_t actually executed depends on a probability measure we introduce below.

In addition, we define the set of all admissible strategies:

$$\mathcal{A}_t = \left\{ S_t : \int_t^T S_\tau^2 d\tau < \infty \right\}, \tag{3.2}$$

that is we assume that all strategies in $\mathcal{A}$ are self-financing and $\mathcal{F}$-measurable with filtration $\mathcal{F}$ defined in Section 3.2.5. We further assume that m and w are bounded from above and δ from below. The self-financing property, the square integrability, and the boundedness are standard

[16] We distinguish between δ as control variable and $l(\delta)$ as the amount of shares actually posted to the market with a mark-up of δ.

assumptions ensuring regularity of optimal controls and are used in the proof of the optimality of the solution. We refer to Federico et al. (2010) for a detailed discussion of these properties.

The utility-maximizing strategy can be visualized as a path between z_0 and $z^* = 0$ through time. Note that we find the optimal strategy using a dynamic programming approach, meaning that we find optimal controls as a function of the current state. This means that the optimal path is only known and can only be plotted *ex-post*. Figure 3.1 depicts the intuition behind this idea.

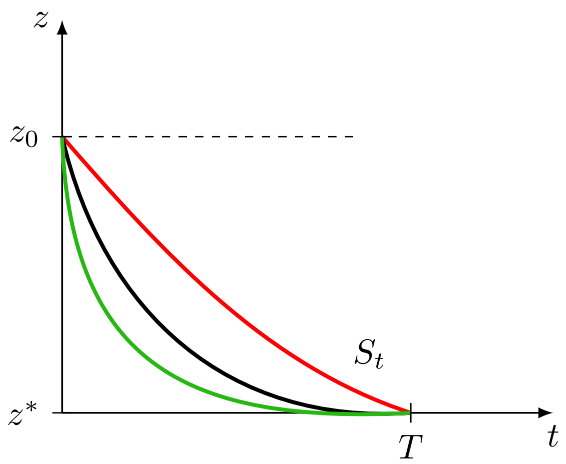


Figure 3.1: Set of *ex-post* liquidation paths through time, which liquidate the initial inventory z_0. Note that these paths can only be plotted ex-post, as they are dynamically generated by the optimal liquidation policy.

3.2.2 Price Process

In our basic model (Section 3.2), the share mid-price $P \equiv (P_t)_{0 \leq t \leq T}$, is determined from an initial condition, $P_0 = P$ and follows:

$$dP_t = \sigma \mathrm{d}W_t, \tag{3.3}$$

where W_t is the standard Brownian motion. This is the most basic assumption on the share price process, as known from the classic Merton problem (Zhu and Ma, 2018). We make this choice of the process, because it substantially simplifies the closed-form solution of our basic model.

In our model extension (Section 3.3.3), we let the price process be influenced by the aggregate rate of trading of all market participants. We choose the "mean-field" formulation similar to Cartea and Jaimungal (2016). Mean-field refers to the aggregate of the actions of other market participants, faced by the trader.

Given some linear permanent impact (slope) parameter, $b \geq 0$, and (mid-price) volatility parameter, $\sigma \geq 0$, the mean-field mid-price evolves according to:

$$dP_t = b(\mu_t - m_t)\mathrm{d}t + \sigma \mathrm{d}W_t, \tag{3.4}$$

where μ_t is the aggregate rate of trading of all market participants, m_t is the trader's rate of market orders, and W_t is the standard Brownian motion. The permanent impact b is defined *net* of a trader's own market order rate.

3.2.3 Trading Mechanism

Traders face inventory costs of η for holding unwanted inventory at every point in time t with $Z_t \neq z^*$. We assume this cost to be equal for all traders and to reflect effects such as holding costs, portfolio (mis)allocation costs, and similar cost components.

When placing a market order, the trader faces a temporary price impact $\kappa > 0$, which is proportional to the rate m_t and has no permanent effect. Note that only the *aggregate* of market orders has a permanent price impact $b > 0$ as described above. The linear temporary price impact is a standard assumption ensuring no-arbitrage property (Huberman and Stanzl, 2004). We refer to Curato et al. (2017) for a detailed discussion on temporary and linear price impacts.

In addition, we assume the limit order to have no permanent influence on the price process, as it is filled by a market order which "walks the book". To account for execution uncertainty, we impose a probability on the limit order being filled:

$$\mathbb{P}[fill|\delta] = \frac{2(P_{max} - \delta)}{P_{max}^2} =: \mathcal{P}_\delta, \tag{3.5}$$

where P_{max} is defined as the maximum price at which an order can still be lifted. While in theory, there is no upper bound on the price, P_{max} can be calibrated empirically. We refer to Section 3.4 for a detailed discussion.

For our extended model (Section 3.3.3), we assume a different parametrization: $\mathbb{P}[fill|\delta] = \kappa_\ell(1 - \kappa_\ell\delta/2)$, with $\kappa_\ell > 0$.[17] This formulation ensures that the resulting first-order condition is analytically tractable. For a more detailed discussion of this choice, we refer to Appendix A.2.1.[18]

We further introduce a Poisson process, $\mathcal{N}^\delta$ with intensity λ_δ that, without loss of generality, represents the arrival of buy orders.[19] Not all buy orders are large enough to "walk the book" until the posted limit sell order. We introduce N^δ as the counter (controlled process) for orders sufficiently large to reach the posted limit order. The probability of an increment in $\mathcal{N}^\delta$ resulting in an increment of N^δ is defined by (3.5). We refer to Guéant et al. (2012), Guo et al. (2016) for a more elaborate model of optimal placement in limit order books.

3.2.4 Workup Session Mechanism

The workup sessions always take place at the current mid-price. We allow market participants to submit parts of their undesired inventory to the workup session. We assume that volumes in the size discovery session arrive according to an independent Poisson process, $N^w \equiv (N_t^w)_{0\leq t\leq T}$, with intensity λ_w. We assume that the arrival intensity does not change throughout the trading session and that traders can reasonably approximate it from past LOB data. Let Ξ_t be the cumulative volume available for size discovery at any given time. We assume that $w_t \ll \Xi_t$, i.e., the volume posted to the workup session can always be matched by the "other side of the book". This assumption ensures the analytical tractability of our solution. An obvious extension here would be to impose a probability distribution of the workup volume being filled. We refer to Buti et al. (2017) for a detailed discussion on similar mechanisms.

[17] Note that $P_{max} \equiv \frac{2}{\kappa_\ell}$ recovers the other parametrisation.

[18] Note that the probability of being filled at any (continuous) point in time can be interpreted as the probability of a given (discrete) order being filled (binary).

[19] As we assume that the trader in our model is a net seller, his limit orders are matched against buy orders arriving "on the other side of the book".

3.2.5 Information Flow

All market-related information such as average price response, market volatility, cumulative inventories, as well as intensities λ_δ and λ_w are publicly known. This assumption is similar to the standard Almgren-Chriss model (Almgren and Chriss, 2001). All traders observe these parameters directly and use them to derive their strategy S_t. We assume that our controls u are predictable and càdlàg with respect to a filtration $\mathcal{F} \equiv (\mathcal{F}_t)_{0 \leq t \leq T}$, that is generated by P, N^δ, and N^w. These conditions ensure that the optimal strategy can be computed ahead of time and that all required parameters are known.

3.2.6 Decision Problem of the Trader

The inventory held by a trader at every point in time evolves according to:

$$dZ_t = -m_t dt - l_t dN_t^\delta - w_t dN_t^w. \tag{3.6}$$

Equation (3.6) shows that the change in inventory is influenced by three processes. $m_t dt$ describes the change associated with market orders. $l_t dN_t^\delta$ describes the change associated with limit orders, as we integrate with respect to matching orders arriving on the other side of the book, N_t^δ. Last, $w_t dN_t^w$ describes the change associated with posting volume to the workup session, and it is counted with respect to the workup session's volume, dN_t^w. This inventory results in the following cash process:

$$dC_t = (P_t - \kappa m_t) m_t dt + (P_t + \delta_t) l_t dN_t^\delta + w_t P_t dN_t^w. \tag{3.7}$$

The cash process is a sum of three components: the first component is the volume traded via market orders at the market price, reduced by the temporary (quadratic) price impact factor κ: $(P_t - \kappa m_t)$. The second component is the cashflow from limit orders, represented by the mid-price P_t, adjusted by trader's mark-up δ_t, and multiplied with the volume sold. Similarly, the proceedings from the workup session are represented by the volume traded times the mid-price, as we assume that there is no price impact there.

We impose a quadratic penalty $\alpha \geq 0$ for liquidating the asset and a linear penalty $\eta \geq 0$ for holding the asset during the trading horizon $[0, T]$, as described in Section 3.2.3.

The performance criterion of a trader is to maximize the end-of-horizon utility under risk aversion. This is a well-known performance criterion functional (Frei and Westray, 2015, Cartea and Jaimungal, 2016):

$$J^u(t, C, P, Z) = \mathbb{E}_{t,C,P,Z}\left[C_\tau + Z_\tau(P_\tau - \alpha Z_\tau) - \eta \int_t^\tau (Z_s)^2 ds\right]. \tag{3.8}$$

The criterion is formulated for a stopping time τ as opposed to the end of the planning horizon T. We define τ as the last point in time when $Z_t > 0$. This accounts for cases where the trader liquidates his position prematurely, before the end of the trading session. Intuitively, the performance criterion, therefore, reflects the current cash position, the penalized inventory, valued at the current price (minus the penalty), and a quadratic penalty on the entire path of the inventory process until the end of the trading session.

3.3 Optimal Liquidation Strategies

We first define the value function based on the performance criterion as stated above:

$$J^u(t, C, P, Z) \equiv \sup_{u \in \mathcal{A}} J^u(t, C, P, Z). \tag{3.9}$$

Then we apply the dynamic programming principle to get the corresponding Hamilton-Jacobi-Bellman (HJB) equation to find the optimal strategies.[20]

3.3.1 Derivation

We refer to Appendix A.2.2 for more details on this procedure.

[20] We refer to Appendix A.2.2 for details on the dynamic programming principle and the derivation based on Øksendal and Sulem (2019).

Using the value function (3.9), we solve:

$$
\begin{aligned}
0 = \partial_t J + \frac{1}{2}\sigma^2 \partial_{PP} J - \eta Z^2 &+ \sup_m \{[m(P - km)\partial_C - m\partial_Z]J\} \\
&+ \lambda_\delta \sup_\delta \{\mathcal{P}_\delta [J(t, C + l(P + \delta), P, Z - l) - J]\} \\
&+ \lambda_w \sup_w \{[\mathbb{E}[J(t, C + Pw, P, Z - w) - J]]\},
\end{aligned} \tag{3.10}
$$

with the terminal condition at time T being $J(T, C, P, Z) = C + Z(P - \alpha Z)$. Note that each sup-term represents the value of placing a market order, a limit order of a given depth and the participation in the workup session.[21] Next, we apply the ansatz $J(t, C, P, Z) = C + ZP + j(t, Z)$[22] and plug it into (3.10):

$$
\begin{aligned}
0 = \partial_t j - \eta Z^2 &+ \sup_m \{-\kappa m^2 - m\partial_Z j\} \\
&+ \lambda_\delta \sup_\delta \{\mathcal{P}_\delta [l(\delta)\delta + j(t, z - l(\delta)) - j]\} \\
&+ \lambda_w \sup_w \{\mathbb{E}[j(t, Z - w) - j]\},
\end{aligned} \tag{3.11}
$$

with $j(T, Z) = -\alpha Z^2$.

We now derive the optimal rates of trading for each of the three trading possibilities: market orders, limit orders and workup session participations. We solve each of the terms in (3.11) for m, δ, w, respectively. We apply the ansatz:[23]

$$
j(t, Z) = j_{0t} + Z j_{1t} + Z^2 j_{2t}, \tag{3.12}
$$

and the terminal conditions are transformed to $j_0(T) = 0$, $j_1(T) = 0$ and $j_2(T) = -\alpha$. With these assumptions and ansatz[24], we find the following optimal strategies:

$$
m_t^* = \frac{-j_1 - 2Zj_2}{2\kappa}, \tag{3.13}
$$

[21] Intuitively, the first sup-term represents the change in cash and in inventory resulting from market orders. The second sup-term represents the change in value from placing a limit order with a quantity of l, and a mark-up of δ and the third sup-term represents the effect of posting w shares to the workup session.

[22] The value function, J, is thus decomposed additively into the present (time t) cash from liquidation (in both lit and dark markets and limit orders), C, plus the mid-price (P) marked-to-market book value of the remaining inventory (Z), ZP, plus the *excess* book value, j, given the present inventory.

[23] This is a standard ansatz applied to the performance criterion as defined in 3.8, compare e.g. Cartea and Jaimungal (2016). The excess book value, j, is assumed quadratic in inventory, Z.

[24] In the following, we omit index t for functions j_0, j_1, j_2 for better readability.

$$\delta_t^* = \frac{1}{2}\left(j_1 - aj_2 + 2j_2 Z + P_{max}\right), \tag{3.14}$$

$$w_t^* = \frac{j_1 + 2Zj_2}{2j_2}. \tag{3.15}$$

We refer to Appendix A.2.4 for a detailed derivation of these optimal controls. To find j_0, j_1, j_2 we resubstitute optimal controls into (3.11) and collect all terms for (Z^0, Z, Z^2), i.e. we collect all scalar terms, those dependent on Z linearly and those depending on Z quadratically. We refer to Appendix A.2.5 for a detailed derivation. Solving this system of PIDEs (Appendix A.2.6) yields the optimal strategies m^*, δ^*, w^* in closed form. For better readability, we set $C_{1/2} = \frac{1}{2\theta}\left(\lambda_w \pm \sqrt{\lambda_w^2 + 4\theta\eta}\right)$, $\theta = \left(\frac{1}{\kappa} + \lambda_\delta \frac{a}{P_{max}^2}\right)$, $\Theta = \theta(C_1 - C_2)$ and $\Lambda = \frac{\alpha + C_2}{\alpha + C_1}$. Then, the optimal controls are given by:

$$m_t^* = \frac{-Z}{\kappa} \frac{C_2 - C_1 \Lambda \exp(-\Theta(T-t))}{1 - \Lambda \exp(-\Theta(T-t))}, \tag{3.16}$$

$$\delta_t^* = P_{max} - \left(\frac{a}{2} - Z\right) \frac{C_2 - C_1 \Lambda \exp(-\Theta(T-t))}{1 - \Lambda \exp(-\Theta(T-t))}, \tag{3.17}$$

$$w_t^* = Z. \tag{3.18}$$

3.3.2 Properties of the Optimal Liquidation Strategy

We find that the optimal controls $\{m^*, \delta^*, w^*\}$ all depend on each other: the functional forms of m^* and δ^* contain both, λ_δ and λ_w. Similarly, w^* depends on m^* and δ^* over their respective effects on the inventory process Z_t. Next, the solutions remain defined for a wide range of parameters. The only requirement we have to impose is $\lambda_w^2 + 4\theta\eta \geq 0$. We did not explicitly constrain the controls to be positive. However, due to the martingale property of the price process, all three controls move downwards for a net seller.

Rate of Market Orders

In general, we see that the market order rate is positive and accelerates as T comes closer and the $exp(-\Theta(T-t))$ term converges to 1. We refer to Appendix A.2.7 for the proof. The rate declines with temporary price impact κ. We observe that $\frac{\partial \Theta}{\partial \lambda_w} > 0$.[25] This means that with an increasing amount of orders posted to the workup session, the market orders are posted at a lower rate at any point $t < T$. Similarly, we find that $\frac{\partial \Theta}{\partial \lambda_\delta} > 0$.[26] With an increase in liquidity λ_δ, the amount of buy orders large enough to walk the book until the posted limit sell order increases. This leads to less aggressive bidding with market orders.

On the other hand, we find that the rate becomes faster, as the trader is increasingly penalizing $Z_T \neq 0$, i.e., becoming increasingly risk-averse with respect to unwanted inventory holdings by the end of trading. Taking this to the limit, we find $\lim_{\alpha \to \infty} \Lambda = 1$, therefore making the rate m^* the fastest.

Rate of Mark-ups in Limit Orders

The optimal mark-up for the limit order posted by the trader is expressed as P_{max}, reduced by a factor dependent on all other strategy parameters. Intuitively, this means that as long as the outstanding inventory is larger than the amount traded in a limit order at once (a), the trader keeps on posting limit orders. This follows from $\left(\frac{a}{2} - Z\right) \frac{C_2 - C_1 \Lambda \exp(-\Theta(T-t))}{1 - \Lambda \exp(-\Theta(T-t))} < 0$ as long as $\frac{a}{2} < Z$. This also means that at a point in time close to T, $\mathcal{P}_\delta = 0$ and therefore, no further limit orders get executed. We refer to Appendix A.2.8 for the proof. This result means that when the trader decides to post large-sized limit orders, he will stop trading these earlier. We also observe that the orders are posted more aggressively for higher α, because $\lim_{\alpha \to \infty} \Lambda = 1$ and therefore, the mark-ups are smaller.

Rate of Participation in Workups

The rate of participation in workups is maximized by the trader throughout the entire trading session. This is to be expected, as workups present the least risky way to trade in our model.

[25] $\frac{\partial \Theta}{\partial \lambda_w} = 2\lambda_w \sqrt{\lambda_w^2 + 4\eta \left(\frac{1}{\kappa} + \frac{a\lambda_\delta}{P_{max}}\right)} > 0.$

[26] $\frac{\partial \Theta}{\partial \lambda_\delta} = \frac{4a\eta}{P_{max}} \sqrt{\lambda_w^2 + 4\eta \left(\frac{1}{\kappa} + \frac{a\lambda_\delta}{P_{max}}\right)} > 0.$

This result can be interpreted as the trader always trying to liquidate his residual inventory, net of the market, and limit orders, in the workup session. We find that the rate of arrival of liquidity in the workup session λ_w influences the rate m^* with $\frac{\partial m^*}{\partial \lambda_w} < 0$. This is in line with the theory that workups limit liquidity in the lit market. We numerically investigate the effect of the intensity of the workup session in more detail in Section 3.3.4.

3.3.3 Model Extensions

In the preceding section, we derived the optimal strategy for a single trader acting in the market as defined in Section 3.2. While such a model is suitable to derive qualitative properties of optimal strategies, it does not reflect the interaction between traders. As there is no back-loop between the optimal strategy S_t^* and the price P_t, apart from the temporary price impact κ, we cannot make any statement about a market equilibrium in that model. The extension presented here draws from the mean-field model by Cartea and Jaimungal (2016) and allows us to impose an order flow process on top of the price process. The order flow process represents the cumulative effect of all other traders.

To model the market with order flow, we first introduce two additional, independent Poisson processes $L^{\pm} \equiv (L_t^{\pm})_{0 \le t \le T}$, both with intensity λ_f.[27] These processes represent the order flow on both sides of the order book(bid/ask). We modify the drift process:

$$d\mu_t = -\kappa_f \mu_t dt + \eta_f (dL_t^+ - dL_t^-), \tag{3.19}$$

with $\kappa_f, \eta_f > 0$ being volatility and decay parameters, respectively. As we take the difference between L_t^+ and L_t^-, we are modelling the impact of the net order flow.

We also modify the counting process for the processing of limit orders, N_t^δ:

$$N^\delta \equiv (N_t^\delta \le L_t^+)_{0 \le t \le T}.$$

This ensures that the number of market *buy* orders, L_t^+, which lift the limit *sell* order of the agent/walk the sell side of the book to a price not exceeded by $P_t + \delta_t$.

[27] In the following, we use the subscript f for all parameters related to order *flow*.

With these modifications, the HJB Equation (3.8) is now given by:

$$J(t, C, P, \mu, Z) \equiv \mathbb{E}_{t,C,P,\mu,Z}\left[C_\tau + Z_\tau(P_\tau - \alpha Z_\tau) - \eta_f \int_t^\tau (Z_u)^2 du\right]. \tag{3.20}$$

And the modified resulting HJB (previously 3.10) is given by:

$$\begin{aligned} 0 = \partial_t J + \frac{1}{2}\sigma^2 \partial_{PP} J + \mathcal{L}^\mu J - \eta_f Z^2 + \sup_m\{[m(P - km)\partial_C + b(\mu - m)\partial_P - m\partial_Z]J\} \\ + \lambda_f \sup_\delta\{\kappa_\ell(1 - {\kappa_\ell \delta}/2)[J(t, C + P + \delta, P, \mu, Z - 1) - J]\} \\ + \lambda_d \sup_w\{[J(t, C + Pw, P, \mu, Z - w) - J]\}, \end{aligned} \tag{3.21}$$

with the terminal condition, $J(T, C, P, \mu, Z) = C + Z(P - \alpha Z)$.

The infinitesimal generator of the net order flow is given by:

$$\begin{aligned} \mathcal{L}^\mu J = -\kappa_f \mu \partial_\mu J + \lambda_f[J(t, C, P, \mu + \eta_f, Z) - J] \\ + \lambda_f[J(t, C, P, \mu - \eta_f, Z) - J]. \end{aligned} \tag{3.22}$$

We refer to Appendix A.2.9 for the detailed derivation of optimal strategies $\{m^*, \delta^*, w^*\}$ resulting from this model, as well as for definitions of all parameters. Here, we discuss how order flow changes the strategies. These are given by:

$$m^* = -\frac{1}{2\kappa}\left(b - 2\frac{\beta}{\xi}\right) Z \left[\frac{-1}{2\kappa}(-\lambda_f \kappa_\ell^2 \bar{\ell}_0(t, T; 2\gamma) + b\mu\bar{\ell}_1(T - t) + \; 2Z\chi)\right], \tag{3.23}$$

$$\delta^* = \frac{1}{\kappa_\ell} + \frac{1}{2}\left[(1 - 2Z)\frac{\beta}{\xi} - \lambda_f \kappa_\ell^2 \bar{\ell}_0(t, T; 2\gamma) + b\mu\bar{\ell}_1(T - t) + (2Z - 1)\chi\right], \tag{3.24}$$

$$w^* = Z - \frac{1}{2}\frac{\lambda_f \kappa_\ell^2 \bar{\ell}_0(t, T; 2\gamma) - b\mu\bar{\ell}_1(T - t)}{\chi - \frac{\beta}{\xi}}. \tag{3.25}$$

The optimal trading speed (3.23) differs from that of Almgren and Chriss (2001) by its three leftmost summands. The addition of workups is captured by the second summand, $\bar{\ell}_1(T - t)$. We see that $\bar{\ell}_1(T - t) \xrightarrow{t \to T} 0$ continues to intuitively hold: order flows influence the optimal trading speed of the trader less as maturity nears. Also, note that the share of inventory Z is now adjusted by the term $\left(b - 2\frac{\beta}{\xi}\right)$ which intuitively captures the trader's temporary and the market's permanent price impacts, as well as the workup intensity and the probability of limit orders being filled.

Second, the optimal mark-up (3.24) is now scaled by the term $\frac{\beta}{\xi}$, accounting for price impacts, and directly depends on the market order flow via $\bar{\ell}_1$. Note that the rightmost term of (3.24) corresponds to our base model, as χ has a similar functional form.

Third, we note that the optimal volume w^* posted to the workup session is marked down by a factor mostly influenced by the intensity of the net order flow λ_f and the permanent price impact b faced by traders in the lit market.

Last, we find that $\bar{\ell}_0(t, T; \lambda_d)$ has a complicated influence on all three controls (3.23)-(3.25), when $\lambda_f, \kappa_\ell > 0$. To study the statics of this influence, either a numerical approximation or an assumption such as $\lambda_d = n\gamma$ for $n \in \mathbb{N}$.

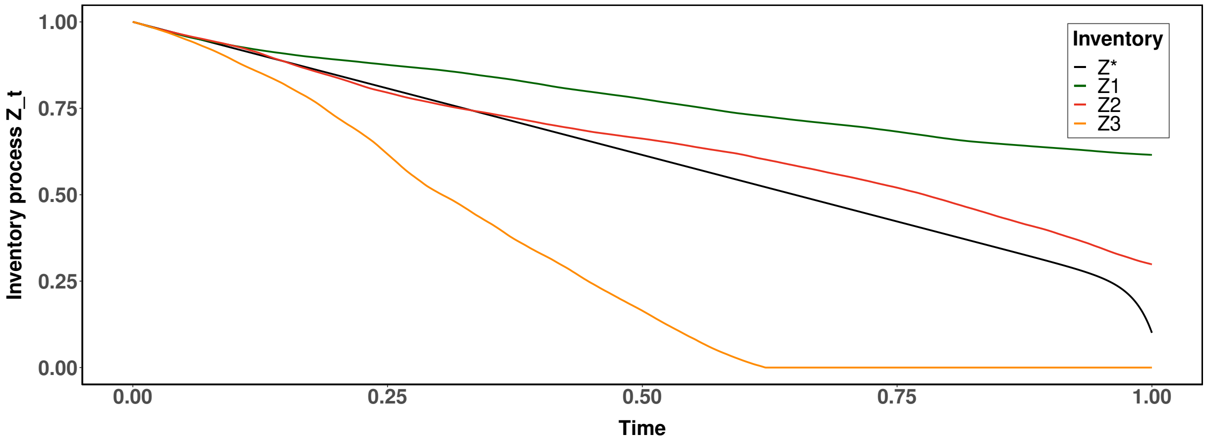
In sum, we note that while the closed-form solution is still available for optimal controls with order flow, numerical simulations might offer a more direct way to study the comparative statics.

3.3.4 Numerical Simulations

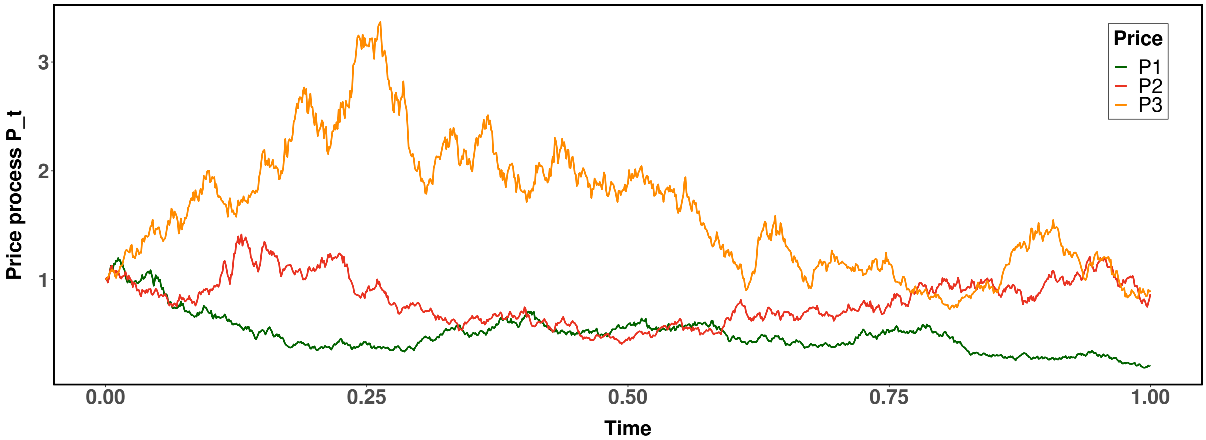
We simulate our main results on optimal liquidation strategies obtained in Section 3.3.1 to shed light on their statics. As the model relies on randomness in prices, arrivals of workups, and fill probabilities of limit orders, we are randomly sampling each of these processes. We then apply a discretized solution of the main HJB Equation (3.10) to these samples to arrive at optimal inventory curves.

We find that the solution of the liquidation strategy is very sensitive with respect to λ_δ and λ_w. For the simulations in Figure 3.2, we set $\alpha = 10, \eta = 50, \lambda_\delta = 2, \lambda_w = 5, \kappa = 1$. The price process is sampled using $\mu = 0.1, \sigma = 0.8$. We sample random prices three times (Scenarios $P1, P2, P3$) and generate three optimal inventory paths ($Z1, Z2, Z3$). We plot Z^*, the theoretical inventory path along with the generated paths.

Under these parameters, we observe that the trade-off between α and η is such that the $Z_T > 0$, i.e., the trader is incentivized to limit the area under the allocation curve more than he is to reach $Z_T = 0$. We also observe that the strategies behave as expected with respect to the price process P_t. In the case of Scenario 3 ($Z3$), the corresponding prices $P3$ are quite high during the first half of the trading session, so it is optimal to liquidate the entire position before T. This also demonstrates that the stopping time formulation works as expected when $\tau < T$. Similarly,



(a) Simulated inventory processes over time



(b) Simulated price processes over time

Figure 3.2: Side-by-side comparison of inventory and price processes, values over time. Part (a) of this figure shows how the inventory develops over time for 3 different scenarios. The black line (Z^*) represents the theoretical optimal inventory as derived from the model. Part (b) shows the corresponding samples of the price processes.

for relatively low prices with little drift, the incentive to liquidate is lower, and the trader reaches a significantly higher residual inventory in Scenario $Z1$.

The optimal liquidation path behaves regularly with respect to both risk aversion parameters α, η. When shifting more weight towards the terminal penalty, we observe that all paths converge to 0 before T, regardless of the price process, as can be seen in Figure 3.3.

We find that in the practical application, a calibration to both, the market price process and the own risk preferences is required to achieve sensible liquidation strategies.

In addition, we consider the empirical average slope of the process Z_t as the speed of reallocation. We find that decreasing λ_w by 1% decreases the speed of reallocation by 0.6%.

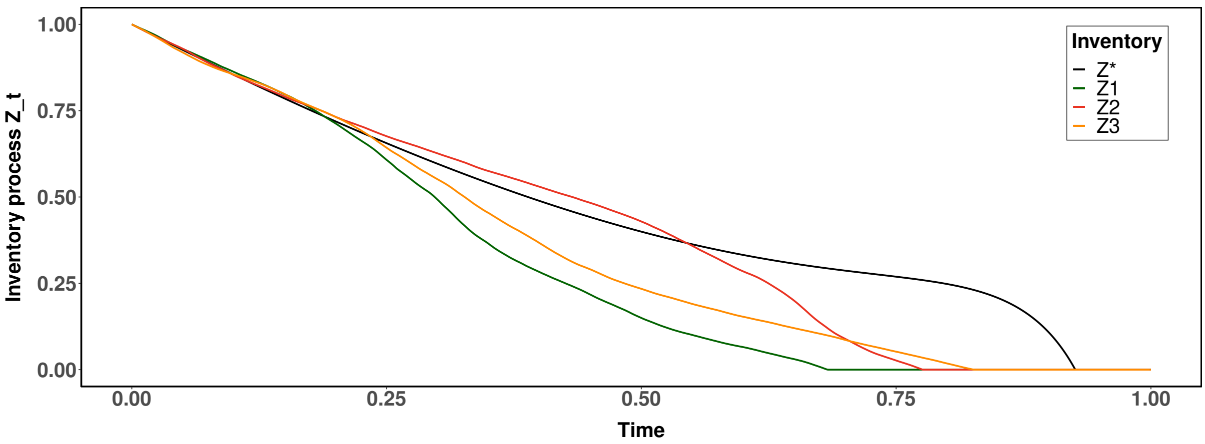


Figure 3.3: this figure shows how the inventory develops over time for 3 different scenarios under high risk aversion parameter α. The black line ($Z1$) represents the theoretical optimal inventory as derived from the model.

3.4 Simulation with Order Book Data

We illustrate our results using real-world LOB data. Our aim is twofold: first, we want to show how our optimal liquidation strategies perform compared to established benchmark strategies. Second, we want to investigate the effect of randomized size-discovery sessions on a real LOB, in particular in terms of the reallocation speed. While the actual effect of introducing a new mechanism can only be tested in practice, our aim in this simulation is to approximate the actual behavior of market participants as closely as possible.

3.4.1 Data

We employ LOB data from the NASDAQ Nordic exchange. Our dataset spans three years, starting from 2018 and ending in 2020. We observe monthly, quarterly, and yearly power futures. NASDAQ is offering two kinds of power futures: the regular power futures are settled mark-to-market on a daily basis, while the deferred settlement futures (DS) are accumulating the mark-to-market valuation during their trading period with a cash settlement coming on the 20th day of the following calendar month. This means that the current mark-to-market is not realized throughout the whole trading period. For monthly contracts, only DS futures are observed in our dataset. As these futures are not being traded on any other exchange, we observe the complete available liquidity of these contracts, excluding possible OTC trading. This is a rare setting in the LOB research literature (Gould et al., 2013).

All futures are quoted in euro, and the contract base size is 1 MWh. Depending on the delivery period, the number of delivery hours for monthly and quarterly contracts can vary between 672–745 hours and 2159–2209 hours, respectively.
We consider baseload contracts only, meaning that delivery takes place every hour of the day (Nasdaq Nordic, contract specification, 2021).

We observe the orders submitted to each contract's LOB as a series of event messages, ordered in the sequence of their submission to the order book. In the case of NASDAQ, the series is generated by the ITCH protocol (Nasdaq Nordic, ITCH protocol, 2021). We observe the complete depth of the order book, i.e., limit orders placed at an arbitrary level, as well as every change of the LOB on an event-by-event basis. This means that every alteration of the order stack, including order submissions, order modifications, and cancelations, is observable. For a summary of event messages as specified by the protocol, we refer to Table 3.1.

Table 3.1: ITCH protocol: message specifications

Field	Values
Message type	Add (A), Delete (D), Execute (E), Revise (R)
Timestamp	Timestamp of order's arrival, measured in nanoseconds
Order book ID	Unique identifier of the order
Symbol	Contract traded: ENOM*/ENOQ*/ENOY*
Quantity	Volume to be transacted
Type	Bid / ask order
Price	Price of the order
Order attributes	Market / Limit / Short sell / Override / Fill-or-kill

Notes: each message of the ITCH protocol consists of these fields. As NASDAQ Nordic is using one protocol for very diverse contracts, not all values can be found in data on power futures. In particular, orders like "fill-or-kill" are not found in our data. We refer to Nasdaq Nordic, ITCH protocol (2021) for further details on the protocol.

3.4.2 Empirical Strategy and LOB Matching

Our empirical strategy for testing the theoretical results is split into two parts. First, we are simulating regular LOB trading with market and limit orders while placing these using the optimal strategy for market and limit orders from Section 3.2. In our notation, this is equivalent to setting $\lambda_w = 0$. Second, we simulate a workup session by performing regular LOB matching and combining it with workup-like matching for a subset of orders in the order book at random times.

In both cases, we evaluate each contract's LOB separately, building up the order queue based on bids and asks submitted only for that contract. The orders placed by each strategy are added to the order book. We then match the order book based on both, the orders already present in the dataset (other participants) and the orders that we placed.

We first introduce the necessary notation. The state of the order book is a set of all bids and asks currently present in the queue. We denote the set of all bids active at time t as $B_t = \{b^{bid}, q^{bid}\}_t$ and the set of all asks as $A_t = \{b^{ask}, q^{ask}\}_t$. We denote the slope of the bid and ask side of the order book by $\psi^{bid/ask}$ where

$$q^{bid/ask} \approx \beta_0 + \psi^{bid/ask} b^{bid/ask}, \tag{3.26}$$

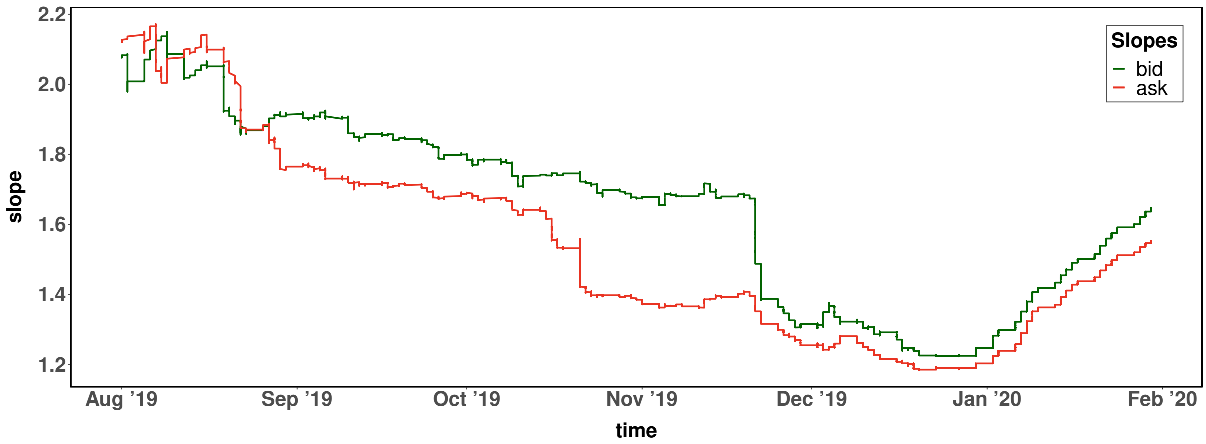
i.e., it is a linear approximation of the slope on both sides of the order book. We define $\psi^{bid/ask} > 0$, regardless of which side of the book is considered. We refer to Figure A.2 for a depiction of this concept. As the sets A_t and B_t change with every tick, we can estimate (3.26) for every tick. This yields an irregularly spaced time series of slopes, which we consider a proxy for the liquidity of the contract over time. Figure 3.4 shows two exemplary time series of slopes.

We assume that the trader holds a "natural" net short position in power. This is usually the case of a power producer. We set $Z_0 = 50$ MWh for each day. This number is chosen such that the temporary price impact of individual orders is significant enough (i.e., the order actually walks the queue of limit orders by more than 1 step), but at the same time, the overall price impact after placing the position is not significant. We measure this by comparing the average bid-ask spread with and without placing an order. We allow for a deviation of up to 5% in bid-ask spreads here.

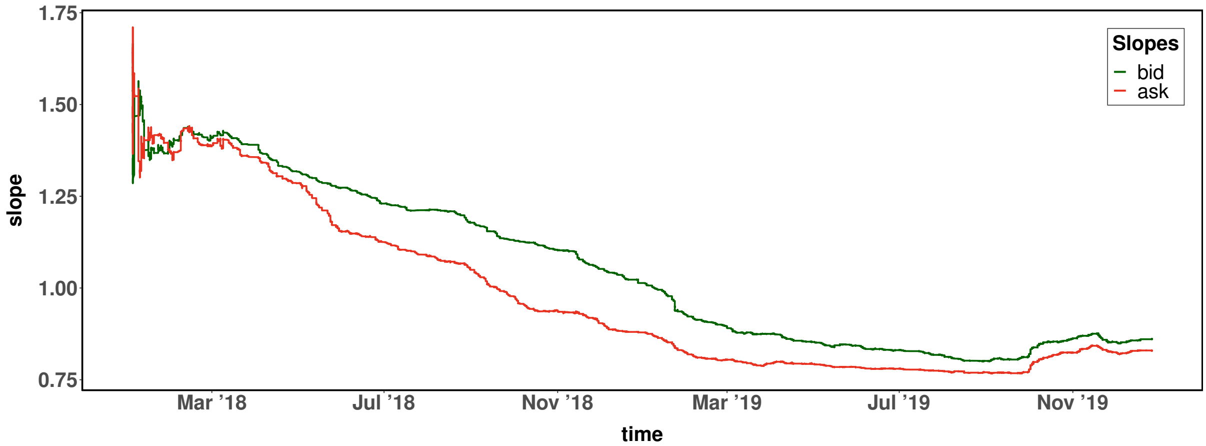
P_{max} is calibrated such that the deviation between the theoretical density of a limit order being filled and the actual probability of it being filled (until the end of the session) is minimal. We refer to Figure 3.5 for a graphical representation of this idea.

In addition, we set $\alpha = \eta = 10$, the variance of the price process σ_t is estimated based on the empirical intraday variance of the prices, computed on a rolling basis for the last 30 days.

The price impact of a market order is denoted by κ. It is estimated by placing market orders (b^{ask}, q^{ask}) with a size ranging from 1MWh to 30MWh. Each placement yields a price impact



(a) Slopes for the monthly base load contract, delivering February 2020



(b) Slopes for the quarterly base load contract, delivering Q1 2020

Figure 3.4: empirical slopes of both sides of the order book for a monthly and a quarterly contract (slope over time). Part (a) of this figure shows how the slopes of the monthly contract developed over time, part (b) shows the same development for a quarterly contract. The shape of the curves is representative for other contracts.

$P_t - \tilde{P} > 0$, as we are placing sell orders. Then,

$$P_t - \tilde{P}_t = \alpha_0 + \kappa q_t^{ask}, \tag{3.27}$$

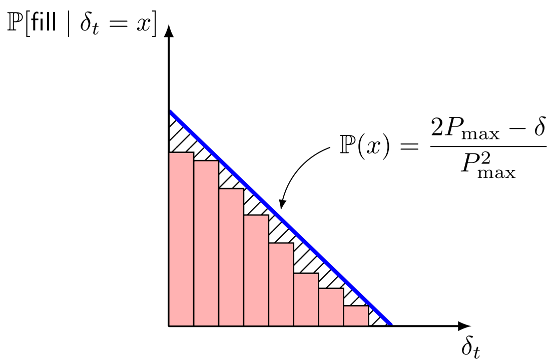


Figure 3.5: Calibrating P_{max} to the actual order fill probability. The histogram represents the order probabilities as measured by placing an order δ above the initial price of the day and waiting until the end of the day (T) to check if it gets filled. The blue line is the density of a triangle distribution fitted to the histogram.

where $\tilde{P}_t$ is the actual price from the transaction, including the effect of "walking the book". The price impact will be different depending on the time the market order is set. To account for this effect, we sample the point in time when we set the order uniformly. When applying this technique, we find an average κ of 0.45.

3.4.3 Continuous Trading

We assume that a trader wants to liquidate a positive position ($Z_0 > 0$). We assume T to be the end of one trading day. We select benchmark strategies to compare the optimized strategy against. As benchmarks, we select established optimal liquidation strategies (Frei and Westray, 2015, Sağlam et al., 2019). These include the time-weighted average price (TWAP), the volume-weighted average price (VWAP), and the liquidity-weighted average price (LWAP). We define $\bar{P} = \frac{1}{Z}\sum_i P_i Z_i$ with Z_i being the volume traded in trade i and Z being the total volume.

Definition 3.1 (TWAP strategy)**.** Let $P_{TWAP}[0,T] = \frac{1}{T}\sum_{k=1}^{N} P_k d_k$, where d_k is the time during which the price P_k remains constant. The time-weighted average price strategy is minimizing $|P_{TWAP} - \bar{P}|$.

Definition 3.2 (VWAP strategy)**.** Let $P_{VWAP}[0,T] = \frac{1}{V}\sum_{k=1}^{N} P_k v_k$ be the volume weighted average price of a trading session, with $V = \sum_k v_k$ being the total volume and v_k the total volume transacted for price P_k. The volume-weighted average price strategy aims to minimize $|P_{VWAP} - \bar{P}|$.

Definition 3.3 (LWAP strategy)**.** Let $P_{LWAP}[0,T] = \frac{1}{L}\sum_{k=1}^{N} P_k l_k$ be the liquidity weighted average price of a trading session, with L being the sum of all slopes of the bid side, ψ^{bid} and P_k the price of a trade. The liquidity-weighted average price strategy aims to trade N units of the contract at an average price that corresponds to P_{LWAP} as closely as possible, i.e. it is minimizing $|P_{LWAP} - \bar{P}|$.

We discretize time in 15-minute intervals to compute these strategies. We implement the TWAP strategy by placing a quantity of $\frac{Z}{28}$ every 15 minutes of the trading day, trading over 7 hours (7 hours $\times$ four 15-minute-intervals = 28 intervals). The VWAP strategy is implemented by taking the volume profile of the preceding day as a proxy for V and 15-minute average volumes of the day before as a proxy for v_k. We place trades proportional to each 15 minutes volume's share of the cumulative volume of the day before. For the LWAP strategy, we assume the slope ψ^{bid}

from the day before is a proxy for L. We place a quantity of $Z\frac{l_k}{L}$ every 15 minutes, where l_k is the average of all slopes belonging to this 15 minutes interval.

3.4.4 Continuous Trading with Workup Sessions

For this simulation, instead of taking the perspective of a trader liquidating a position, we consider the perspective of the exchange settling orders arriving in the order book. We consider order book matching with and without a randomized workup session.

We measure the impact of the workup session arriving with an intensity λ_w in terms of volume transacted per time, average price, price volatility, and average liquidity represented by the bid-ask spread. These are the typical metrics observed by market participants and represent measures of the favourableness of market conditions.

To simulate participation in the workup session by orders recorded in our LOB, we make some assumptions. We know from our model in Section 3.3.3 that the likelihood that a trader participates in the workup session depends on the execution risk of his limit orders, the time pressure he is facing, and the volume of the order. To account for these factors, we make the following assumptions:

Assumption 1. At any given point in time, $t \in [0, T]$, top $\gamma_{wu}\%$ of open orders in terms of volume always participate in the workup session. These are assumed to be traders facing the largest potential market impact.

Assumption 2. The likelihood of participation in a workup session for any limit order in the book increases by 1/7 for each hour after 09:00, until 6/7 for the last trading hour of the day.

Assumption 3. The execution risk for natural sellers is higher than for natural buyers. All else equal, an open limit order to sell participates in the workup session with a probability of 1 and a limit order to buy with a probability of $buy^{wu} = 0.9$.

Assumption 4. The workup session arrival intensity λ[28] is calibrated such that, on average, at least one session takes place every hour, i.e., we set $\lambda = 1$, and we sample inter-arrival time at the beginning of the trading day from the $exp(\lambda)$-distribution.

[28]Note that this is conceptually a different λ, i.e., not λ_w.

In sum, our assumptions ensure that participation in a workup session is sufficiently high. We need these assumptions, because it is not clear that in a Nash equilibrium, all traders would participate in workup sessions.

3.4.5 Results

Continuous trading without workups

In this section, we report the results of applying the optimal strategy from Section 3.3 to the LOB. We find that the optimal liquidation strategy S_t^* achieves prices and price volatilities very similar to those achieved by the LWAP strategy. This is to be expected, as the optimization problem (3.10) is optimizing the speed of liquidation vs. the costs of liquidation. As it takes the price impact explicitly into account, its results are similar to trading during the most liquid times, which is, in turn, similar to weighting the volume traded by the current expected price impact. In addition, we find that the deviation between volume-weighted and liquidity-weighted prices can be substantial. A possible interpretation here is that traders transacting large volumes do not always do so when market liquidity is optimal. Alternatively, the fact that large volumes are being transacted could be the reason that the market liquidity is not optimal. In both cases, it is evident that a liquidity-oriented strategy can lower the variance of prices while reducing adverse effects in settlement prices.

Table 3.2: Continuous trading without workups: market metrics

	TWAP		*VWAP*		*LWAP*		S_t^*	
	Avg. price	σ	Avg. price	σ	Avg. price	σ	Avg. price	σ
2018								
Monthly	45.59	9.88	46.15	10.18	47.39	11.11	47.14	9.80
Quarterly	38.80	9.44	38.78	10.01	39.05	10.87	39.02	9.15
Yearly	34.90	5.80	34.25	6.28	35.00	6.00	35.79	8.96
2019								
Monthly	42.13	6.55	43.22	6.73	42.77	6.98	42.07	6.26
Quarterly	36.55	6.61	37.61	6.72	37.26	7.01	37.88	6.80
Yearly	30.85	3.81	32.16	4.02	31.07	4.54	31.00	3.98
2020								
Monthly	21.10	12.90	23.15	13.71	23.02	12.85	23.16	11.50
Quarterly	29.65	7.07	28.03	7.74	30.02	8.01	31.07	7.15
Yearly	29.80	2.95	29.78	2.87	30.16	2.55	30.65	2.70

Notes: All metrics are averaged over the entire lifetime of a contract. We cluster all contracts by their type and average over all contracts of the same type. The months August and September 2019 contain missing observations and are not considered in this sample.

Continuous trading with workups

We report our results in terms of average volume transacted, average price, volatility, and bid-ask spread. Average volume refers to the volume average in MWh transacted per 1 hour of trading during the trading session, averaged over all contracts in our sample. We consider this a measure of allocative efficiency. Average prices indicate how much a workup session influences the price process. Price volatility is the average intra-hourly volatility of the price. From the perspective of market participants with physical positions in power markets, this is a measure of risk. Last, the bid-ask spread is taken as a measure of liquidity.

We find that workup sessions positively influence the transacted volume with an increase in the transacted volume of up to 16%. The effect is especially pronounced for monthly contracts. This result follows directly from the fact that we are matching more limit orders at any point in time. The impact on the average price is insignificant, meaning that neither the demand nor the supply side pays for the speedup in reallocation by getting a significantly higher (or lower) price. We see, however, that the prices are slightly increasing. This is most likely explained by the ranking of liquidity on the bid and the ask side. We find that for most contracts, the ask

Table 3.3: Continuous trading with workups: market metrics

	Metric	*Trading mechanisms* CT	WU 1	WU 2	WU3	WU 4
Monthly	Avg. volume	43.01	45.20	45.35	45.02	49.95
	Avg. price	34.25	34.98	34.50	35.01	34.30
	σ	9.95	9.84	9.85	9.11	9.15
	BAS	0.27	0.35	0.49	0.47	0.43
Quarterly	Avg. volume	58.20	66.28	66.33	65.01	66.98
	Avg. price	35.09	35.01	35.12	35.59	34.72
	σ	7.95	7.99	7.18	7.24	7.63
	BAS	0.22	0.30	0.34	0.31	0.35
Yearly	Avg. volume	46.01	51.15	51.47	51.99	51.85
	Avg. price	31.74	31.89	31.58	31.77	32.12
	σ	4.20	4.34	4.11	4.30	4.35
	BAS	0.15	0.20	0.21	0.25	0.23
buy^{wu}		–	0.90	0.90	0.95	1.00
γ^{wu}		–	0.05	0.10	0.10	0.10

Notes: All metrics are averaged over the entire lifetime of a contract. σ represents price volatility as described above. We cluster all contracts by their type and average over all contracts of the same type. The months August and September 2019 contain missing observations and are taken out of our test.

side exhibits flatter slopes, meaning that the ask side is less price sensitive. This asymmetry between the bid and the ask-side liquidity is a commonly observed phenomenon (Cenesizoglu and Grass, 2018, Sensoy, 2019). For that reason, the amount of participants in a workup session for any given moment is higher on the bid side. This, in turn, might be driving the prices up. We observe a slight decrease in volatility, which is to be expected because market participants are transacting higher volumes in a workup session without influencing the price. Last, the average bid-ask spreads increase, as workup sessions take liquidity out of the market. This is again in line with extant studies on the impact of dark pools (Buti et al., 2017). We found this effect to not be very robust and to depend a lot on the parametrization of the workup session. In terms of how fast an order book recovers to the old state of liquidity after a workup session, we consider again the slopes of the bid and ask side and check how fast these revert to their old levels after a size-discovery session. Figure 3.6 displays an exemplary development of ψ^{bid}. We find that the recovery in slopes is observable, but the slopes typically do not recover to their old, pre-workup level.

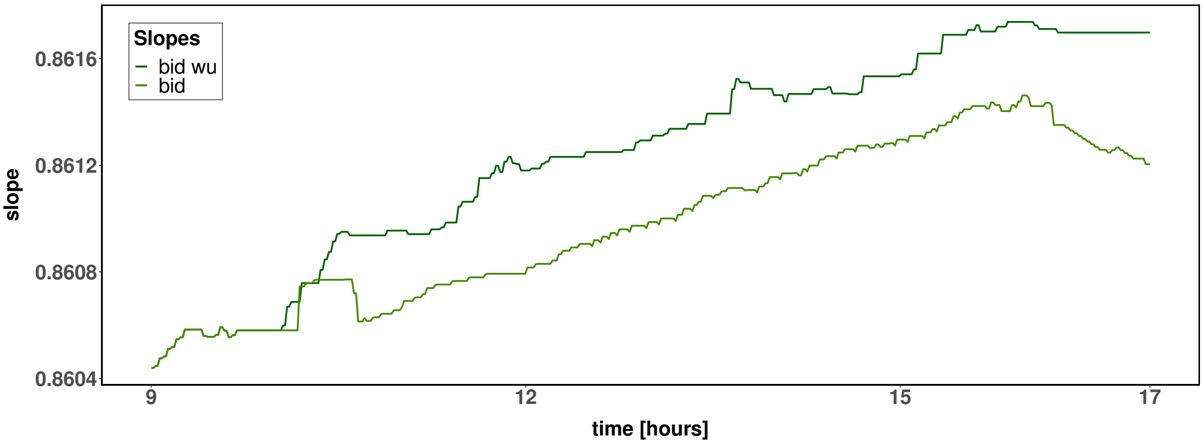


Figure 3.6: ψ^{bid} development over time for the trading session on 27th of December, 2019. The effect of the workup session can be seen in the difference between the two slope time series. Overall, the workup sessions bring up the price sensitivity of the bid side of the order book. Towards the end of the session, workups stabilize the liquidity. This can be explained by the majority of orders already being filled.

3.5 Conclusion

We have introduced a market model which combines important features of the optimal liquidation problem, such as continuous time execution, various order types, size discovery sessions, risk aversion, temporary and permanent price impacts, and order flow.

We have used this model to show that workup sessions speed up the reallocation, at the expense of reducing liquidity supplied to the market via limit orders. On the other hand, we have found that traders following the optimal liquidation strategy reduce their market orders, thereby consuming less liquidity. In a numerical setting, we have illustrated that the inventory paths generated by our model are effective in liquidating the undesired inventory when parametrized with the right risk preference. Empirically, we have shown that following the optimal strategy yields less variance in settlement prices. We have also argued that introducing a random workup to an existing LOB can speed up the reallocation.

Nevertheless, practical issues arise when adapting such a mechanism: First, traders' strategies are extremely sensitive with respect to the random volume arriving in a workup session. Second, price impact and liquidity calculations become significantly more complex, resulting in higher complexity in product pricing and risk management. Third, price discovery might be impaired. On the other hand, adopting such a regime would potentially result in the reunification of continuous markets and OTC trading in dark pools, thus bringing the overall liquidity available on the markets up. It is a classical result that higher liquidity results in more efficient allocations.

4 | Fairness vs. Welfare: Disclosure in Financial Markets

Vadim Gorski, Philipp Maume, Martin Kellner[29]

The reporting of corporate insiders' trading (Directors' Dealings, DD) is a cornerstone of regulatory efforts to foster market efficiency. However, little is known about its effectiveness. We investigate regulatory effectiveness based on an event study methodology, adjusted to our purposes. We employ recent data on DD from Germany, the UK, and the USA. We argue that the DD regulations are not as effective as shown in previous studies, especially in the case of the UK and the USA. Our findings suggest that an indicator effect is achieved at the expense of fairness. Further, we find that trade volume, insider level, and publication period of DD influence regulation effectiveness. In addition, we provide empirical evidence that the regulatory framework is poorly enforced. Based on these insights, we derive policy implications.

[29] Author contributions: This essay is based on a joint paper with co-authors. My contribution was the design of the empirical strategy, the empirical work and the draft of major parts of the paper.

4.1 Introduction

Company directors, officers, and other employees are insiders enjoying the possibility to trade company's stocks. This gives them an informational advantage over other market participants, as they might have access to private knowledge about the state of their company. This poses a problem, because insiders utilize this advantage to achieve excess returns of up to five percent over 20 days after the trade, as shown by numerous studies (Seyhun, 1986, Lakonishok and Lee, 2001, Friederich et al., 2002, Hillier and Marshall, 2002, Gregory et al., 2009, Dymke, 2011).

This calls for regulatory measures to control insiders' trading[30] via disclosure. The reasoning is that if insiders disclose their positions, the information becomes public and can be priced in. However, the exact implementation of the disclosure regime poses a challenge. On the one hand, insiders fulfill an important signaling role when conducting their trades. Disclosing these too far in advance would potentially harm insiders and reduce their incentives to trade.[31] On the other hand, fairness of market outcomes is an important aim, too. If the insiders are not obliged to disclose their trades, they are more likely to reach excess returns. An effective disclosure policy has to balance the aim of signaling (indicator effect) with the aim of fairness.

The ability to measure the signaling effect (in the following referred to as the indicator effect) and the fairness is a key prerequisite for effective regulation. A common statistical approach to measure the effect of insiders' trades are event studies (Corrado, 1989, MacKinlay, 1997). Event studies aim to measure the excess return induced by the trade compared to a "normal", counterfactual return. These methods, however, exhibit drawbacks when it comes to assessing insider trading. The first issue is the correct compounding of returns. As McLean (2012) shows, compounding can introduce a substantial bias in event studies. Not compounding, however, introduces a (downward) bias, too. Second, the studies mostly consider the impact of the event by taking only the pre-event or only the post-event time into account. In the case of insiders' trades, both the pre-event time and the post-event time play a role.[32] Last, and more generally, the event studies typically establish a link between the event and the excess return, but not between regulations in a given country and their effect on the fulfillment of regulatory aims.

[30] The corporate officers' transactions in the scope of disclosure regulation (hereinafter: "insiders' tradings") should not be confused with the illegal trading of public shares based on material, nonpublic information about the company (hereinafter: "insider trading").

[31] For example, a large sale transaction by an insider could result in a downward spiral, as other market participants react to this signal.

[32] Pre-event time plays a role in judging the optimality of the point in time when an insider decides to trade. Post-event time plays a role in judging how insiders' position develops.

In our study, we address these two issues: the balance of two opposing regulatory aims, and the drawbacks of event studies and corresponding statistical methods. In particular, we conduct a comparative study over three of the largest financial markets (Germany, UK, USA) and derive the magnitude of excess returns of insiders (indicator effect), as well as their imitators (fairness). Our aim is to enable regulators to make a sensible choice of appropriate measures based on these findings. In addition, we are the first, to our knowledge, to consider excess returns between the day the trade takes place, and the day it is disclosed. We use this measure to determine actual returns achieved by corporate officers between the trade and its disclosure. This provides an important metric for regulators aiming to choose the right policy on the maximum duration between the trade and its disclosure.

We address the statistical drawbacks by providing the following methodological enhancements to the current event study approach: first, we explicitly model compounding returns of each trade. Apart from receiving a better estimator for the holding positions' excess performance, this helps to ameliorate the issue of rebalancing bias as described by Lyon et al. (1999), McLean (2012). Second, we introduce a timing estimator. The timing estimator takes the pre-event returns, and the post-event returns into account and measures the optimality of a given day to perform a DD transaction. For example, a well-timed purchase transaction takes place after a downward trend, right before it shifts to an upward trend. For inference and hypothesis testing, we employ both, parametric and more recent non-parametric statistical tests, as introduced by Kolari and Pynnonen (2011). We moreover perform a set of binary regressions to establish a link between trade volume, insider level in the company, publication period, and abnormal returns.

Our data comprises Directors' Dealings(DD) from July 2005 to September 2015 as reported to the corresponding authorities in Germany, the UK, and the USA. The dataset is compiled by 2iQ, an external data provider specializing in the reporting of insiders' trades. The information on insiders' trades is enriched with metadata on companies and insiders.

We show that abnormal returns of insiders vary between 0.4% and 0.8% in all three financial markets. We also find that German insiders achieve particularly good timing when trading their stock: their transactions take place at ex-post optimal points in time more often. Our results also demonstrate that a longer time until the publication of the DD transaction is negatively affecting fairness in the German market. In contrast, for the UK and the USA markets, our results show that more fairness is promoted by a longer publication period. Lastly, the insider-level variable is positively correlated with excess returns. In particular, executive-level insiders of German companies earn an additional 1.2% excess return within ten days after the transaction as opposed to their non-executive-level colleagues. This applies similarly to UK and USA with 0.7% and 1.7%, accordingly. We, therefore, suggest regulating the ability to purchase own stock, especially for executive insiders, in order to achieve a higher level of fairness.

We relate to the extant literature on DD disclosure in several ways. Empirical studies by Hillier and Marshall (2002), Bajo and Petracci (2006), Dymke (2011), Lee et al. (2011), King et al. (2015), Hodgson et al. (2020) find that insiders are able to generate excess returns. This is in line with our results. However, these studies focus on just one market, whereas we study excess returns among different financial markets with different disclosure regulations.

Dardas and Güttler (2011) investigate European stock markets and compare the magnitude of the announcement effect (indicator effect) among them. Similar to our study, they link these returns to the properties of the transaction, such as insider level or volume. However, they do not compute the returns actually achieved by insiders and the link between these returns and the returns of their imitators.

Some studies find that insiders do not achieve abnormal returns. Friederich et al. (2002) study abnormal returns on the London Stock Exchange. They find that after adjusting for transaction costs, the abnormal returns become insignificant. Dickgiesser and Kaserer (2010) come to a similar conclusion for the German market. We address the problem of transaction costs by only considering the most liquid securities in each of the markets. These securities exhibit lower bid-ask spreads, thereby reducing bias in our estimates.

We also relate to studies distinguishing between excess returns of insiders when buying or selling shares. Lakonishok and Lee (2001), Tebroke and Wollin (2005), Dardas and Güttler (2011), King et al. (2015) find that the sale of shares generates lower to non-existent excess return, as opposed to share purchases. In our study, the highest excess returns are also generated by purchases of shares. However, we still find a significant effect for parts of sales transactions, too.

Gregory et al. (2009), Knewtson et al. (2010) study the timing of trades by corporate insiders. They show that insiders are contrarian investors in the sense that they exhibit a particularly good timing for their transactions. In our study, we introduce a metric for timing consistent with abnormal returns and confirm the hypothesis that insiders exhibit particularly good timing when placing their trades compared to the general market.

We also relate to the information hierarchy hypothesis established by Seyhun (1986) who provided empirical evidence that in the US market, insiders with higher ranks, such as chairmen, achieve higher returns. Our analysis confirms these results, but also relates excess returns to other properties of the DD transaction, such as volume and the period between publication and the trade.

Goncharov et al. (2013) study the effects of governance on excess returns of insiders. We relate to this discussion by showing the role of regulatory enforcement on the effectiveness of disclosure policies.

The remainder of the article is organized as follows: in Section 4.2, we lay out the theoretical foundations for DD regulation and introduce the main purposes of DD regulation. In Section 4.3 we derive statistical properties of CAAR and apply this new metric to our data. We also introduce a timing indicator. In Section 4.4, we present our results. Section 4.5 discusses regulatory shortcomings which might explain our results and derive policy implications. Section 4.6 concludes.

4.2 Director's Dealings Regulation

4.2.1 Regulatory Purposes

DD regulation primarily focuses on the promotion of fairness and efficiency through transparency. There are three main regulatory aims. First, the prevention and detection of prohibited insider trading; second, facilitation of an indicator effect for other market participants (mostly non-insiders) through disclosure of DD transactions; third, promotion of fairness by limiting the excess returns that insiders can generate in comparison to other market participants via timely disclosure.[33] We concentrate on the indicator effect and fairness goals.

Indicator Effect

It is widely agreed that DD transactions are an indicator for the development of a company, as well as for the development of related financial instruments.[34] DD disclosure informs the markets

[33]For regulatory goals of securities regulation as defined by the International Organization of Securities Commissions, we refer to International Organization of Securities Commissions (2010).
[34]Regulation (EU) No 596/2014 on market abuse (hereinafter: MAR).

about sales or purchases made by a corporate officer. Assuming that the officer is rationally motivated, the disclosed transaction could give other market participants an indication about the issuer's wellbeing in the future. By disclosing information relevant for the valuation of shares, the information gap between insiders and other potential investors narrows.

However, DD transactions are not always a reliable indicator. Insiders can have various reasons for buying or selling shares. In particular, the sale of shares can be triggered by sudden cash-flow requirements or by tax considerations. Another trigger can be liquidity needs or diversification purposes (Dymke and Walter, 2008), as well as strategic behaviour (Rahman et al., 2020). Another possible explanation for the systematic difference of reliability of the indicator effect between buy and sell transactions argued about is the higher litigation risk associated with DD sales. If an insider buys stock withholding private information (i.e., the stock is undervalued), the harm caused by this transaction is limited, since it represents only a missed opportunity to generate returns. In contrast, if an insider sells stock withholding private information (i.e., the stock is overvalued), then investors holding a long position in the stock are actually harmed, as the stock prices might fall, thus providing grounds for litigation (Alldredge and Cicero, 2015).

But despite these uncertainties, the indicator effect is not just a theoretical deliberation. In the US, specialized service providers have been established to focus on the collection and analysis of DD disclosure statements, achieving excess returns of three to five percent by trading on the DD disclosure information. (Fried, 1997)

Fairness

Fairness is one of the cornerstones of the integrity of securities markets. The aim is to create a "level playing field" for market participants. The theory of market integrity stipulates that more investors are willing to invest in fair markets (Seligman et al., 2011) and that markets without equally informed market participants will deteriorate in the long run (Akerlof, 1978). The overall economic effect of strengthened market fairness is increased investor confidence, resulting in a higher demand for financial instruments, and ultimately in lower capital costs. The most common way of promoting confidence is transparency (MAR, Rec. 58).

The idea that insiders' superior knowledge allows them to generate excess returns is a stark contradiction to the idea of a level playing field. This tension is the main reason why – despite theoretical deliberations about the merits of permitting insider trading in general (Manne, 1966)

– insider trading is prohibited. Consequently, DD regulation also seeks to promote fairness by limiting the excess returns that corporate officers can generate. Even the mere perception of unfairness can tamper with the trust of market participants in a well-functioning market. For this reason, regulation should not just try to limit unfairness, but also try to avoid the appearance of unfairness.[35]

4.2.2 Regulatory Approaches

To achieve the main goals of promoting fairness and the indicator effect, regulators employ three distinct approaches.

First, DD transactions can be disclosed after they have been executed (so-called "post-trading disclosure"). This "disclose or abstain" approach relies on the idea of market efficiency in its semi-strong form (Fama et al., 1969). [36] The challenge for policymakers here is to find the right time period until disclosure. For example, if the disclosure occurs one week after the transaction, the indicator effect and perceived market fairness will be low. This means that only timely disclosure can be an effective regulatory tool (Fleischer, 2002, Weiler and Tollkühn, 2002). As of today, the vast majority of DD regulation relies on post-trading disclosure.[37]

Second, regulation could require disclosure before a transaction is executed (Fried, 1997).[38] Pre-trading disclosure can achieve a higher level of fairness, because corporate officers can only generate low excess returns. Similar to post-trading disclosure, the disclosure period needs to be kept short. If the corporate officer needed to inform the markets too much in advance, the market price could adapt before the transaction is carried out, and the insider would suffer financial losses.

A third option and ultima ratio is to prohibit DD transactions entirely. While this option maximizes the fairness effect, it harms the indicator effect. It would also harm young companies,

[35] See, e.g., U.K. Model Code, Introduction, Listing Rule 9, Annex 1: "... ensure that persons discharging managerial responsibilities do not abuse, and do not place themselves under *suspicion of abusing*, inside information that they may be thought to have...".

[36] Popular examples are sec 16(a) of the US Securities Exchange Act of 1934, the new Art. 19 MAR, as well as its predecessor, Art. 6(4) of the Market Abuse Directive (No. 2003/6/EC, hereinafter "MAD").

[37] The only exceptions are Rule 9.2.8 (UK) and Rule 144 (US). The first comprises a duty to request clearance to deal from the company and to complete the deal within two trading days of clearance being received. The second rule requires pre-trading registration with the SEC for the sale of certain restricted securities.

[38] A system resembling pre-trading disclosure can be found in Rule 144 regarding the Securities Act of 1933 (17 C.F.R § 230.144, 1994), which requires pre-trading registration with the SEC for the sale of certain restricted securities (Coffee Jr et al., 2015).

which are often owned by the directors, who are usually also the founders. Temporary bans represent a less severe alternative.[39] One commonly used example is the restricted period of 30 days before the quarterly reporting date.[40] To sum up, the trade-off between the indicator effect and the fairness can be found by the right combination of disclosure periods (i.e., periods between the transaction and its disclosure), as well as prohibition of insiders' trades around critical dates.

4.3 Data and Empirical Strategy

We consider the excess returns of the corporate officer and of the imitating trader (imitator). The excess returns of the imitator constitute the indicator effect, and the excess returns of the corporate officer constitute the fairness effect. This allows us to measure fairness and indication separately.

Our focus lies on three important capital markets: Germany, the UK, and the USA. The basis of our empirical analysis is a dataset obtained by 2iQ Research, a data provider that gathers commercial information from various sources such as supervisory authorities, stock exchanges, and economic information services. 2iQ Research continuously checks the quality of the data acquired. We cross-check the German and the US part of the data with the data obtained directly from regulatory authorities.[41] It should be mentioned that especially the regulatory data sources can suffer from poor data quality.[42]

We filter the dataset to only include securities of the German, British, and the US-American benchmark indices (DAX, FTSE100, DJIA). We delete all trades that took place before July 1, 2005, the date the Market Abuse Directive (MAD) was to be implemented by the EU member states (Ventoruzzo, 2015). These deletions result in a sample of trades in most liquid stocks that took place in a stable regulatory environment.

[39] These bans are also called "closed periods" or "restricted periods".

[40] As imposed, for example, in the UK Listing Rule 9.2.8. Similar regulations can be found in Listing Rules 12.9 – 12.12 of the Australian Securities Exchange.

[41] In the case of Germany, we received an extract of DD transactions directly from BaFin.

[42] A prominent example for the poor data quality is the so-called Satterwhite stunt: 61 invented trades by the fictitious person "Johnny Earl Satterwhite" were published in the American EDGAR system in the year 2011. The reported shares on Microsoft alone would have been 1.500 times those of Bill Gates, and the number of shares would have exceeded the existing amount of Microsoft shares by a factor of 118.

Overall, our data set consists of DD transactions exercised from July 2005 to September 2015 for German, the UK, and the US-American companies with high market capitalization. For an overview of the sample and sample statistics, we refer to Table 4.1.

Table 4.1: Characteristics of the obtained 2iQ dataset.

	Germany		UK		USA	
Number/% of trans. types	*N*	[%]	*N*	[%]	*N*	[%]
Buy	1,070	41 %	3,482	13 %	865	1 %
Sell	973	37 %	11,128	41 %	43,113	74 %
Award	91	3 %	8,047	30 %	6,923	12 %
Subscription	22	1 %	394	1 %	0	0 %
Derivative	457	17 %	4,128	15 %	7,391	13 %
Transaction volume [TEUR]	μ	*Mdn*	μ	*Mdn*	μ	*Mdn*
Buy	2,989	57	22,658	41	1,156	81
Sell	6,481	648	6,790	227	878	101
Award						
Subscription	23,411	58	6,161	23	0	0
Derivative	873	197	405	81	894	275

Notes: All DD transactions from July 2005 until September 2015 are reported for securities belonging to DAX, FTSE100 and DJIA indices. N is the number of transactions in the sample. The transaction types are classified as follows: *Buy* means a share purchase on the open capital market, *sell* means a share disposal on the open capital market, *award* means a share-compensation for employees, *subscription* means subscribing to a capital increase, *derivative* means the exercise of derivatives (e.g. options, share rights, conversion rights, employee stock options etc.), the transaction volume for each trade is calculated as the product of the actually paid price and the number of shares traded. μ denotes the volume's mean and Mdn its median. The skewness of the distribution of the transaction volumes is due to transactions with exceptionally high volumes as traded by corporate officers via private fonds or family offices.

We observe a relatively small percentage of sell transactions in Germany in comparison to the UK and the USA. An explanation could be that in Germany, stock-based compensation is less common than in the Anglo-Saxon financial markets (Kaserer and Moldenhauer, 2008). Surprisingly (but in line with the aforementioned explanation), in Germany, the most common transaction type is a share purchase. In contrast, the US-American stock purchases account for only one percent of reported DD transactions.[43]

We now introduce definitions used throughout our study.

[43] A possible reason is a duty for handing over short-swing profits according to Sec. 16(b) Securities Exchange Act from 1934, requiring the insider to keep the stock for at least six months if profits are to be retained. A plausible consequence would be a lower attractiveness level of own stock purchases.

Event An event is defined as either the insider's transaction or the imitator's transaction. The imitator's transaction event takes place the moment the insider's transaction is being made public, as the imitation takes place by a trader observing the publication of the DD transaction and placing the same trade.

Estimation period and event window The estimation period is used to compute expected (regular) market returns. These are then used to derive abnormal returns. The estimation period usually starts up to 110 days prior to the event and ends at the latest 20 trading days before the event to avoid interference with pre-event price movements (Freyaldenhoven et al., 2019). This results in an estimation window of 90 days. The abnormal returns are then measured between two points in time, t_1, t_2, chosen within the event window, see Figure 4.1.

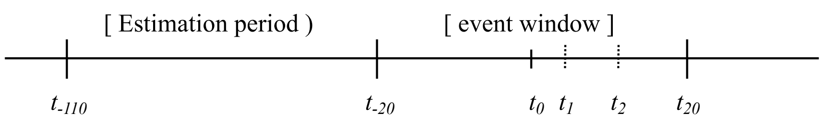


Figure 4.1: Timeline for the event analysis (days relative to event days). Estimation period comprises the days t_{-110} to t_{-21} and the event window comprises the days t_{-20} to t_{20}.

Expected Market Return In line with MacKinlay (1997), expected returns of asset i are defined as

$$\mathbb{E}[R_{i,t}] = \alpha_i + \beta_i R_{m,t}, \tag{4.1}$$

with $R_{m,t}$ being the market return, as observed using an appropriate market index at time t. We chose CDAX for the German sample, FTSE ALL-SHARE for the UK sample and S&P500 for the US sample. The choice of the market indices is based on both, size and relevance considerations. We estimate the parameters α_i and β_i by performing an ordinary least squares (OLS) regression between the returns' time series of security i and its corresponding index. For OLS estimation, we select a 90 days estimation period as described above. This procedure is repeated for each considered event of every security i.

Abnormal Return The abnormal return of asset i at time t (trading day relative to event day) is the difference between actual return and normal (or expected) return, i.e.

$$AR_{i,t} = R_{i,t} - \mathbb{E}[R_{i,t}]. \tag{4.2}$$

Average Abnormal Return (AAR) The AAR for trades $j \in N$ at time t (relative to the event day) is defined as

$$AAR_t = \frac{1}{N}\sum_{j=1}^{N} AR_{j,t}. \tag{4.3}$$

Cumulative Average Abnormal Return ($CAAR$) We define

$$CAAR[t_1, t_2] = \frac{1}{N}\sum_{j=1}^{N}\left(\prod_{t=t_1}^{t_2}(1 + AR_{j,t}) - 1\right), \tag{4.4}$$

i.e. we compute the compounded part of abnormal return caused by the event for every trade j and *then* average over all j's. This expression is commonly approximated by a pure sum over $AR_{j,t}$ in the literature (McLean, 2012). We find that the difference between the approximation and the true value (considering compounding) can be substantial in several cases. Based on our sample, the difference for some trades accounted for up to 16% for $CAAR[1, 10]$. We thus choose to compound only the abnormal part of investors' returns. This is equivalent to measuring the direct effect of excess returns by virtually buying the abnormal return derivative and holding it. Thereby, we consider compounding of abnormal returns, but we do not distort our excess return measure with compounded normal returns.[44]

Timing Estimator (TE) We define the timing estimator as

$$TE[t] = (CAAR[1, t] - CAAR[(-t + 1), 0]) * 100. \tag{4.5}$$

The timing estimator is an additional measure based on the $CAAR$. This estimator relates the excess returns before and after the regarded event. This relation serves as a measure of the impact of the event on the returns. By considering $CAAR$ before and after the event, the underlying timing in relation to the market can be measured. Timing refers to the extent to

[44] We refer to Table A.1 in the appendix for an exemplary comparison between BHAR, CAR w/o compounding and CAR.

which a given transaction meets the peak (or trough) of the share price in relation to the market movement, i.e., good timing means to buy at the lowest point in a given window or to sell at the highest. The interpretation is as follows: High TE for a purchase transaction indicates high negative abnormal return before the purchase and high positive return after the purchase. If the slope of the $CAAR$ does not change within a symmetric event window, TE is equal to zero. The observations of insiders' purchase transactions usually show negative abnormal returns before the purchase trade and positive abnormal returns after the purchase trade. The opposite applies to sell transactions. The timing estimator is usually positive for purchases and negative for sales.

Last, we refer to the Appendix A.3.1 for definitions of statistical tests employed in this study and their properties.

4.4 Hypotheses and Results

4.4.1 Hypotheses

We introduce two groups of hypotheses. The first group of hypotheses tests the overall effectiveness of regulation. We measure the effectiveness using $CAAR$ and TE as metrics.

The second group of hypotheses evaluates the effect of DD transactions' properties (publication period, transaction volume, insider level) on regulation effectiveness. We utilize $CAAR$ as our metric and combine it with binary regression modeling.

Regulatory Effectiveness

We differentiate between two main regulatory goals and reflect them as sets of hypotheses. Hypotheses belonging to the set ***H1*** reflect the goal of fairness, hypotheses from the set ***H2*** reflect the goal of indication.

To test the regulatory goal of fairness, we use the insiders' trading day as the event day and compute excess returns and timing indicators.

To test for the indicator effect, we perform the same procedure, but choose the date of disclosure as the event day, thereby measuring the excess return of an imitator.

H1 Insiders do not achieve abnormal returns or particularly good timing with their trades, i.e., the goal of fairness is achieved, as no unfair advantage of DD transactions is present.

H2 Imitators of insiders' DD transactions achieve abnormal returns or particularly good timing, i.e., the goal of indication is achieved.

Note that ***H1*** and ***H2*** are stated with respect to the desired regulatory outcome. For H1 *no* excess returns is the desired outcome, while for ***H2***, it is the *opposite*: imitators achieving excess returns means that the indicator effect is working. The hypotheses ***H1*** and ***H2*** are further divided into sub-hypotheses pertaining to transaction type, as summarized in Table 4.2.

Table 4.2: Hypotheses and corresponding sub-hypotheses for DD regulation effectiveness

Hypothesis	*H1* (fairness)				*H2* (indication)			
Event	trading day				publication day			
Trans. type	buy		sell		buy		sell	
Sub-hypoth.	H1.1	H1.3	H1.2	H1.4	H2.1	H2.3	H2.2	H2.4
Accept if	$CAAR \leq 0$	$TE \leq 0$	$CAAR \geq 0$	$TE \geq 0$	$CAAR > 0$	$TE > 0$	$CAAR < 0$	$TE < 0$

Confirmation of any sub-hypotheses indicates that the main hypotheses might be true. For example, H1.1 is true, if $CAAR$ is not positive for buy transactions on a trading day, and similarly, H1.2 is true, if $CAAR$ is not negative for sell transactions on trading day.

DD Properties and their Effect on Regulatory Effectiveness

The second group of our hypotheses is employed to investigate whether certain properties of DD transactions affect the achievement of fairness and indication goals. We test these hypotheses using binary regression. The hypotheses are formulated along two dimensions: The first dimension represents regulatory goals, fairness, and indication. The second dimension represents the properties of a particular DD transaction: publication period, transaction volume, and insider level. These are the properties we observe in our dataset. To summarize, our framework is set up according to Table 4.3.

Table 4.3: Hypotheses for possible determinants of DD regulation effectiveness

	Properties of DD transactions					
Regulatory goal	Pub. period (a)	*sign*	Volume (b)	*sign*	Insider level (c)	*sign*
Fairness (1)	H1a	+	H1b	+	H1c	+
Indication (2)	H2a	−	H2b	−	H2c	−

Notes: The sign of the binary regression coefficient needed to affirm the hypothesis is given in the columns denoted with "*sign*".

Each of the six hypotheses is verified using $CAAR$ as a metric for abnormal returns. Note that each of the hypotheses can be evaluated for both, buy and sell transactions. However, due to inherent differences between buy and sell transactions, we only consider buy transactions to verify our hypotheses. As explained above, sell transactions might have a different motivation than just trading the stock.

Publication Period Hypotheses A shorter publication period, meaning the time between an insider trade and its publication, reduces the information asymmetry of insiders and outsiders with respect to time. H1a is an indicator of fairness, and H2a represents the indicator effect.[45]

H1a $CAAR$ is lower for insiders that trade in short publication periods.

H2a $CAAR$ is higher for imitators of trades with short publication periods.

The variable *Publication period* is defined as:

$$\textit{Publication period} = \begin{cases} 1 & \text{, if publication period} \geq 2 \text{ days} \\ 0 & \text{, if publication period} < 2 \text{ days.} \end{cases} \tag{4.6}$$

Transaction Volume Hypotheses The second possible determinant for the DD regulation effectiveness is the transaction volume. De minimis thresholds under which insiders are not obliged to report their trades are mainly established for two reasons: first, small transactions are of limited significance for market participants, and second, the surveillance of many small transactions incurs disproportionate effort. We examine whether smaller transactions result in more fairness (H1b) and whether smaller transactions trigger a higher indicator effect (H2b):

[45] We select two days as a threshold, because it subdivides our observations into two about equal groups.

H1b $CAAR$ is lower for low-volume insider trades.

H2b $CAAR$ is higher for imitators of low-volume insider trades.

The variable *Volume* is defined as:

$$Volume = \begin{cases} 1 & \text{, if transaction volume} \geq x_{0.25} \text{ quantile} \\ 0 & \text{, if transaction volume} < x_{0.25} \text{ quantile.} \end{cases} \tag{4.7}$$

The $x_{0.25}$-quantile represents the upper bound of the bottom 25% of all DD buy transactions' volumes on the corresponding market. We made this choice based on the distribution of volumes. As most observations contain relatively small volumes, the 25% quantile allows distinguishing between "regular", small transactions, and large purchases/sales. Our results are robust with respect to this choice, in that slightly smaller and larger quantiles yield similar results.

Insider Level Hypotheses We test whether the goals of DD regulation are better achieved for executive insiders' trades or non-executive insiders' trades (information hierarchy concept as described in Seyhun (1986)). If the two hypotheses can be confirmed, the goals of fairness (H1c) and indicator effect (H2c) are better achieved for trades of non-executive insiders than for trades of executive insiders:

H1c $CAAR$ is lower for non-executive insiders.

H2c $CAAR$ is higher for imitators of non-executive insiders' trades.

The variable *Insider level* is defined as:

$$\textit{Insider level} = \begin{cases} 1 & \text{, if trading corporate officer is an executive-level insider} \\ 0 & \text{, if trading corporate officer is not an executive-level insider.} \end{cases} \tag{4.8}$$

4.4.2 Results on Fairness and Indication

Figures 4.2 and 4.3 present our main results on $CAARs$ attained by insiders graphically. We observe that within the event window, the curve progression is as follows: negative slope before and positive slope after purchases, meaning insiders are contrarian investors. For sales, we find the opposite behavior, with less prominent slope changes. This is in line with the theoretical reasons why purchases and sales differ, as presented in Section 4.1. In particular, sales transactions can be motivated by other reasons than making profit. It also shows that insiders are able to time the market entry better than an average investor.

Table 4.4 summarizes our results for the ***H1*** hypothesis. We consider $CAAR[1, 10]$, i.e., the returns are measured between days 1 and 10 relative to the event day. The purchase transactions of corporate officers of UK companies show the highest $CAAR[1, 10]$ with 0.84%, similar to the US-American ones with 0.89%. In contrast, German insiders achieve only 0.40%. These findings indicate that insiders generate significant excess returns with their DD transactions. This applies to all investigated markets. In terms of timing, we are able to confirm the Hypothesis H1.3 meaning that the timing of the buy transactions outperforms the market. For sales, we cannot confirm or reject our hypothesis. While the Hypothesis H1.2 is significant statistically, the economic significance here is much lower compared to H1.1, meaning that the outperformance of insiders is more prominent for purchases.

Interestingly, we find that $CAAR[1, Disclosure]$[46] is positive and significant for the US sales, meaning US-American insiders miss excess returns within the time from their sell transaction until its disclosure. This is a further indication for "other-than-return-motives" of DD sell transactions. In addition, by calculating $CAAR[1, Disclosure]$, we demonstrate that the insiders' excess returns do not only exist due to the outsiders' imitation of insiders, but are based on the insiders' private information. This applies to both, the UK and the US purchase transactions.

Table 4.5 summarizes our results on the performance of imitators (***H2***). Our findings here are similar to those on the performance of insiders (***H1***).[47] We find that the purchase transactions of imitators of the US-American companies show the highest $CAAR[1, 10]$ with 0.90%, similar to the UK with 0.81%. In contrast, German insiders achieve only 0.23% and the result here is not significant.

This clarifies that, considering timing and relative returns, an indicator effect correlates positively with a lack of fairness through insiders' timing and profit benefits. Therefore, the indicator effect

[46]We calculate and report $CAAR[1, Disclosure]$ and $CAAR[1, 10]$. The latter metric is used to affirm or reject the stated hypotheses, while the former one is employed to account for returns achieved until trade publication. $CAAR[1, Disclosure]$, therefore, serves as an additional direct proof of an existing unfairness for DD transactions, because it reflects the excess return generated by corporate officers without the signaling effect of the corresponding transaction.

[47]Note that ***H1*** is stated *negatively* (absence of excess returns is a confirmation of ***H1***), while ***H2*** is stated positively (excess returns are a confirmation of ***H2***).

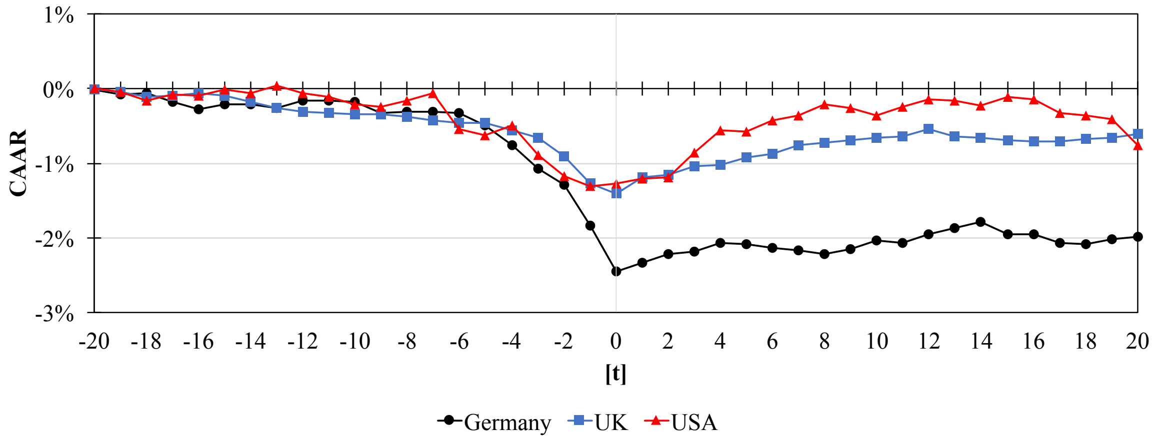


Figure 4.2: Purchase transactions CAAR, with transaction day as event day.

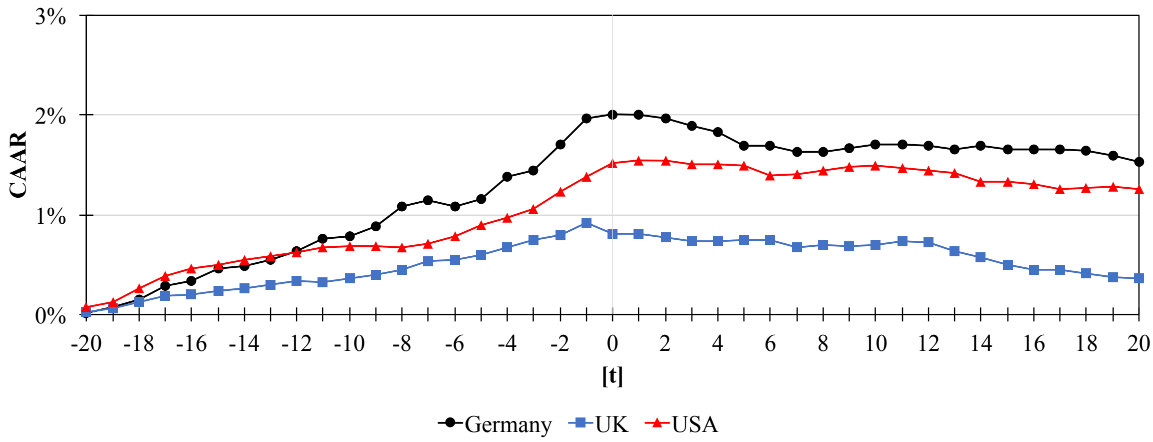


Figure 4.3: Sales transactions CAAR, with transaction day as event day.

requires insiders to outperform the market, and we confirm that the two regulatory goals oppose each other.

In line with this observation, Figures 4.4 and 4.5 show that the change of the slope of the curves is less sharp than for insiders and does not occur at t_0, but slightly earlier around t_{-2}. This suggests that not the publication itself causes the price movements, but the value of information that the corporate officers can materialize by trading their companies' shares.[48]

We also consider the timing exhibited by corporate officers. As can be seen in Table 4.4, corporate officers not only perform well but also do this by finding an optimal point in time for their trades, showing that insiders and imitators are contrarian investors. The best timing is achieved for German DD transactions as insiders have a $TE[5]$ of 2.3. Intuitively, this means that there is a clear, positive change in trend within five days before and after the event. The TE of

[48]See also the results for $CAAR$[1,Disclosure] in table 4.4).

Table 4.4: Excess returns, timing estimators and hypotheses testing results for ***H1***. In this table, we use the 2iQ data as described in Table 4.1.

		$CAAR$			TE	Affirmation of hypotheses?			
		$[-5,0]$	[1,Disclosure]	[1,10]	[5]	*H1.1*	*H1.2*	*H1.3*	*H1.4*
GER	Buy	$-2.16\%^{***'''}$	-0.17%	$0.40\%^{***'''}$	2.3^{***+}	no	no	no	n/a
	Sell	$0.91\%^{***'''}$	$0.05\%^{'''}$	$-0.31\%^{***''}$	-1.2^{***}				
UK	Buy	$-0.92\%^{***'''}$	$0.65\%^{'''}$	$0.89\%^{***'''}$	1.5^{***+}	no	no	no	n/a
	Sell	$0.27\%^{***'''}$	$-0.02\%^{'}$	$-0.13\%^{***''}$	-0.3^{***}				
USA	Buy	$-0.65\%^{***'''}$	$0.85\%^{***'''}$	$0.84\%^{***'''}$	1.2^{***+}	no	no	no	n/a
	Sell	$0.74\%^{***'''}$	$0.35\%^{***'''}$	$-0.02\%^{***''}$	-0.7^{***}				

Notes: In this table, $CAAR[t_1, t_2]$ is defined as in Section 4.3. We set the event date to the trading day on which the DD transaction takes place (e.g. $CAAR[-5, 0]$ stands for the period from 5 days before the trade until the day of the trade). The $CAAR$[1,Disclosure] is the average of all CARs from the day after the transaction until the corresponding disclosure day of that transaction. $^{***/**/*}$ identify significant excess returns and timing estimators based on the 1%/5%/10%-level of the one-tailed parametric t-test. $'''/''/'$ identify the significant excess returns for the GRANK-test. $^{+++/++/+}$ identify significant timing estimators on the 1%/5%/10%-level of the bootstrapped test as described in the methodology section. The hypotheses are answered considering the $CAAR$[1,10] with the GRANK-test and the TE[5] with the bootstrapped test. In the "Affirmation of hypotheses" block, "n/a" stands for cases where we are not able to either affirm or reject the hypothesis.

imitators, on the other hand, is not statistically significant. The German DDs are followed in timing performance by the UK and the US DD transactions.

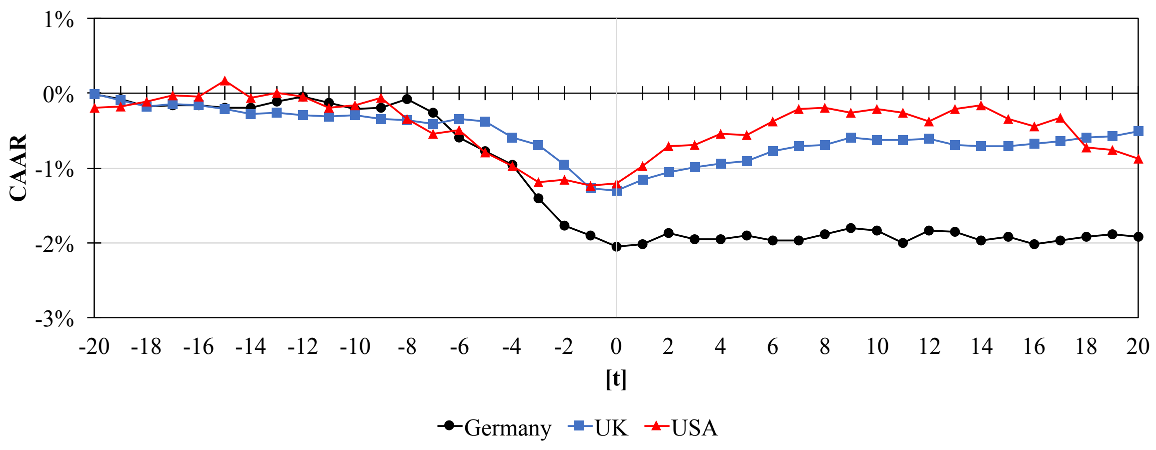


Figure 4.4: Purchase transactions CAAR, with publication day as event day.

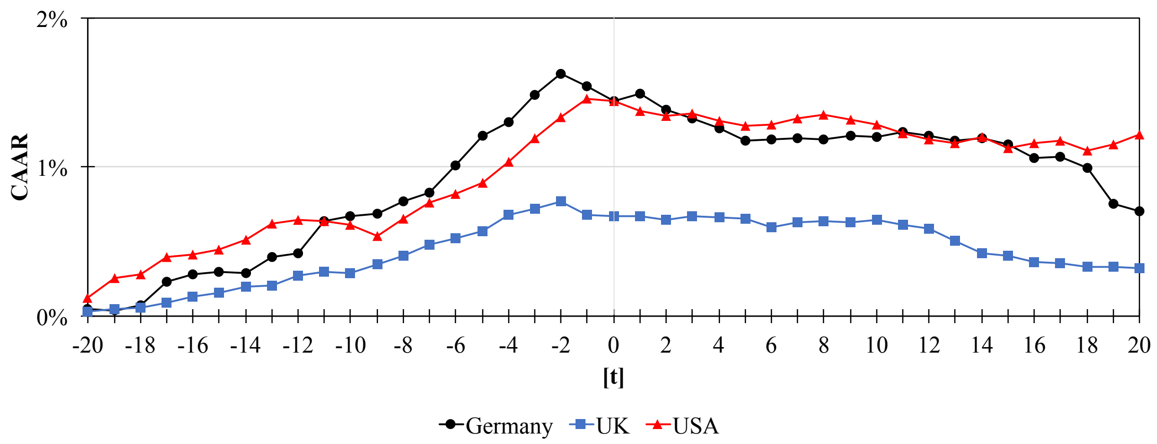


Figure 4.5: Sales transactions CAAR, with publication day as event day.

Table 4.5: Excess returns, timing estimators and hypotheses testing results for ***H2***. In this table, we use the 2iQ data as described in Table 4.1.

		$CAAR$		TE	Affirmation of hypotheses?			
		$[-5,0]$	$[1,10]$	$[5]$	*H2.1*	*H2.2*	*H2.3*	*H2.4*
GER	Buy	$-1.57\%^{***///}$	$0.23\%^{*/}$	$1.4\ ^{***}$	yes	yes	n/a	n/a
	Sell	$0.42\%^{***///}$	$-0.25\%^{*///}$	$-0.5\ ^{***}$				
UK	Buy	$-0.94\%^{***///}$	$0.81\%^{***///}$	$1.4\ ^{***+}$	yes	yes	yes	n/a
	Sell	$0.15\%^{***//}$	$-0.03\%^{***//}$	$-0.1\ ^{***}$				
USA	Buy	$-0.67\%^{***///}$	$0.90\%^{***///}$	$1.0\ ^{***+}$	yes	yes	yes	n/a
	Sell	$0.62\%^{***///}$	$-0.17\%^{***///}$	$-0.7\ ^{***}$				

Notes: In this table, $CAAR[t_1, t_2]$ is defined as in Section 4.3. We set the event date to the publication day on which the DD transaction is disclosed (e.g., $CAAR[-5, 0]$ stands for the period from 5 days before the disclosure until the day of the disclosure). The $CAAR$[1,Disclosure] is the average of all CAR from the day after the transaction until the particular disclosure day. $^{***/**/*}$ identify significant excess returns and timing estimators based on the 1%/5%/10%-level of the one-tailed parametric t-test. $^{///}/^{//}/^{/}$ identify the significant excess returns for the GRANK-test. $^{+++/++/+}$ identify significant timing estimators on the 1%/5%/10%-level of the bootstrapped test as described in the methodology section. The hypotheses are answered considering the $CAAR$[1,10] with the GRANK-test and the TE[5] with the bootstrapped test. In the "Affirmation of hypotheses" block, "n/a" stands for cases where we are not able to either affirm or reject the hypothesis.

4.4.3 Results on Properties of Transactions

We now turn to our results on DD properties and how these influence excess returns. We refer to the Appendix A.3.3 for tables with complete regression results.

Our results for the possible determinants of excess returns and timing advantages indicate that a longer time until the publication of the DD transaction is negatively affecting fairness in the German market. In contrast, for the UK and the USA markets, our results show that fairness is promoted by a longer publication period. Therefore, we obtain mixed results for the hypothesis H1a and cannot establish a definitive link between the time to disclosure and excess returns. This might be because the excess returns of corporate officers are mainly driven by the timing of the trade itself, as opposed to the time until the trade's disclosure.

Pertaining trade volume (H1b) as a determinant of excess returns, our results are mixed, too. Smaller deals exhibit less fairness in Germany, while it is the other way round in the UK. For the US-American market, there is no significant connection between trade volume and additional excess returns.

Lastly, the insider level variable (H1c) is positively correlated with excess returns. In particular, executive-level insiders of German companies earn an additional 1.2% excess return within ten days after the transaction as opposed to their non-executive-level colleagues. This difference becomes even more severe when considering $CAAR[1, 20]$ as a dependant variable: within 20 days, executive-level insiders generate an additional return of 3.3% compared to non-executive-level ones. The coefficients of the insider-level variable related to DD transactions of the USA and the UK securities are also significant. In the case of the UK market, the economic significance of the coefficient is not high.

Our analysis of the determinants of excess returns of imitators shows a similar picture. In terms of publication period (H2a), traders imitating DD transactions tend to perform best by imitating DDs with a shorter publication period. This applies in particular to DD transactions of the US and UK companies. In Germany, the publication does not have a significant impact on imitators' returns.

When considering the trade volume as a possible determinant of excess returns (H2b), the results once again depend on the market: imitators of insiders' transactions in the USA and the

UK perform best by mimicking the trades with a higher transaction volume. In Germany, no connection between trade volume and insiders' imitators' excess returns is observed.

For the insider level (H2c), our results show that the $CAAR$s of imitators of executive-level DD transactions of German and the US-American companies achieve significantly higher returns as compared to the non-executive ones. The UK is an exception with no significant connection between insider-level and imitators' performance here. In summary, transactions made by executive insiders exhibit a particularly good indication.

4.5 Policy Suggestions

In this section, we give an overview of the implications of our findings for optimal regulation. We also shed light on enforcement as a possible reason for differences in excess returns across countries.

4.5.1 Proposed Changes in Regulations

First, our results suggest that only purchase DD transactions exhibit a strong effect in terms of $CAAR$ and in terms of market timing TE. The sale transactions, in general, provide less valuable information to the market and do not generate the same excess returns for the corporate officers. The fact that sell transactions do not allow for economically significant excess returns of insiders and imitators question the necessity of reporting such trades.

Furthermore, we examined DD transactions' properties as possible determinants for achieving the regulatory goals. We found that a higher insider level leads to higher excess returns for insiders and imitators. Our results indicate that if policymakers and regulators focus on fairness, they should restrict DD transactions of executives in particular. In order to ensure an informed market, the disclosure could take place at the time of the transaction (continuous disclosure).

The transaction volume as a possible determinant of fairness and indicator effect leads to conflicting statements. If the regulatory focus was on providing an indicator effect, the reporting threshold for the UK and the US trades could be raised to achieve additional valuable information for each DD on average.[49]

[49] This would imply a detrimental effect on fairness, especially for the UK trades.

For the UK and the US market, we found that trades with a shorter publication period provide better indication than trades with a longer publication period. These results do not apply to Germany. This implies that if regulators choose a longer reporting period, the mitigation of the lack of fairness by the indicator effect becomes weaker. Consequently, the delay between trade and disclosure should be as small as possible to ensure a maximum indicator effect and possibly complete mitigation of the lack of fairness. If the new EU regulatory requirement to disclose not later than three trading days after the transaction (Art. 19(1) MAR) is enforced properly, this would be a first step in the right direction.

4.5.2 Regulatory Effectiveness

We want to put our findings on excess returns into perspective by considering the enforcement of disclosure policies, in this case, disclosure obligations by corporate insiders.

The lack of enforcement of the DD regime is a known issue (Maume and Kellner, 2017). There is evidence that the lack of enforcement might be rendering otherwise effective policies useless (Bhattacharya and Daouk, 2002, Beny, 2005, Bris, 2005). A good regulation regime that is not enforced properly can even have worse consequences than a lack of regulation, as discussed, e.g., by Bhattacharya and Daouk (2009).

In the context of DD transactions, the offender (the corporate officer breaching the disclosure obligation) benefits in two ways. First, the omission to disclose the transaction saves time and money. Second, the offender retains an information advantage over the market participants who comply with the disclosure obligation, allowing him to reap unfair rewards (excess returns). Similar to illegal insider trading, breaches of DD obligations are considered victimless crimes (Manne, 1984). As a result, there is hardly any private enforcement.

Data published by BaFin further supports our claim that disclosure policies are not properly enforced.[50] Table A.4 demonstrates that hardly any fines were imposed in the years 2005 to 2015. A higher number of fines can only be observed in 2006 and 2007, most likely relating to investigations which had been opened between 2002 and 2004.

[50] Other major European regulators such as the French AMF or the Italian CONSOB do not disclose their actions against DD regulation infringements at all.

Second, even if fines were imposed, they were strikingly low. The average fine over the period 2005 to 2015 was 12,700 EUR (a fine of up to 100,000 EUR would have been possible).[51]

Third, the decreasing number of notifications is peculiar. Between 2005 and 2015, the number of notifications dropped from 5,100 to 1,800. Of course, it is possible that officers simply did not engage in DD transactions as often as they had done in the past. However, it is unclear why this should be the case. Moreover, the number of equity trades in the regulated market of Frankfurter Wertpapierbörse (FWB) significantly increased between 2005 and 2015, see Table A.5. In light of this development, the decreasing number of notifications is even more peculiar. To support this finding, we refer to Table A.6.

4.6 Conclusion

We argue that while DD regulation primarily seeks to promote fairness and market efficiency via the indicator effect, these are two opposing goals. A sensible trade-off needs to be found by selecting the right form of disclosure for DD transactions. In any case, promoting market efficiency comes at the expense of fairness.

We provided extensions to the standard event study methodology by introducing a timing estimator, deriving its statistical properties, and considering compounding in the calculation of excess returns. Applying our methodology, we showed that DD transactions provide valuable information for the markets. We demonstrated that insiders achieve excess returns due to their information advantage.

The US-American purchases showed the highest performance for insiders and their imitators, with similar results for the UK and German trades. We affirmed the information hierarchy hypothesis that a higher insider level leads to higher excess returns for insiders and their imitators. Evaluating the effect of transaction volume and publication period on fairness and indicator effect, we come to different results in the three analyzed markets. We question one-size-fits-all approaches and suggest a less incremental approach towards DD regulation.

[51] For comparison: in September 2014 the US Securities and Exchange Commission (SEC) announced in a press release that it had opened investigations against 34 companies and officers for alleged breaches of sec 16(a) Securities Exchange Act. 33 of the alleged offenders agreed to settlements of 2.6m USD in total, resulting in an average of approximately 80,000 USD per offender.

We provided possible reasons for the existence of relatively high excess returns. The problem of effective enforcement is often overlooked. Based on enforcement statistics provided by the German BaFin, we illustrate that there is an obvious lack of enforcement in one of the EU's main capital markets. We conjecture that this is a potential reason for the persisting excess returns generated by insiders. Shorter reporting periods and increased fine limits are important steps towards a fairer and more efficient reporting regime. It may be popular among policymakers to increase enforcement by increasing maximum sentences and fines because it is cost-free. In contrast, increasing funding to enable the enforcement agencies to exercise their functions properly is less popular.

To sum up, our results suggest a less incremental and more strongly enforced approach towards DD transactions and plead for either a pre-trading disclosure or for an instantaneous reporting mechanism, similar to continuous disclosure.

5 | Conclusion: Market Design and its Implications

Traders, exchanges and regulators all aim to better understand market design and which implications it has on various aspects of market activity. For traders, this understanding is mainly tied to the question of how to best implement their desired position in various markets under uncertainty. Exchanges form a market themselves, competing for volume and traders. One of the key competitive advantages in this market are mechanisms that attract traders or make transactions more efficient. Regulators are pursuing different targets still, looking for ways to increase welfare on both sides of the market and trying to avoid strategic behavior of market participants, which harms producers or consumers. For regulators in electricity markets, these questions are even more complex, as, on the one hand, electricity is the key commodity for a functioning economy, and on the other hand, it bears an inherent complexity making the technical and the financial aspects of its delivery inextricable. This thesis sheds light on various aspects of market design and how it affects these three parties. In this chapter, I provide an overview of the key results and their implications.

In the second chapter of this thesis, I focus on measuring demand elasticity in electricity wholesale markets. My essay demonstrates how demand elasticity can be measured using instrument variables in the presence of endogeneity concerns. Such concerns arise when the equilibrium quantity is subject to simultaneous supply and demand shocks. This problem makes it difficult to correctly estimate demand elasticity, which is an important factor in a multitude of electricity market applications, including fundamental market models, policy evaluations, and optimal trading. I investigate how machine learning approaches (Bajari et al., 2015, Abadie and Kasy, 2019) can be applied to improve elasticity estimates. In electricity markets, the true submitted demand curves are known and can be used as a benchmark for model estimates.

This provides a setting for comparing different estimators. In particular, I argue that, in the presence of many and possibly unfit instruments, applying regularization in the first stage of the estimation problem can help to select the right instrumental variables, thereby making the estimates more precise. I demonstrate this effect empirically based on a dataset from the European Energy Exchange (EEX) day-ahead market. I show how, over time, fuel prices become less relevant as supply shifters, while renewable feed-in forecasts remain relevant throughout. My results have several implications. First, they should be taken into account by researchers designing econometric models with instrumental variables. Second, regulators interested in the true demand elasticity of a market can apply these results to estimate it more precisely. Third, market participants can design their pricing strategy based on these estimates.

Apart from measuring market outcomes, another key problem faced by exchanges and regulators is designing market mechanisms. For traders, the problem is to optimally act within these mechanisms. In the third chapter, I focus on workup sessions and, in particular, size discovery sessions (Duffie and Zhu, 2017). Size discovery sessions represent a mechanism to address strategic bid shading in continuous LOB markets. Strategic bidding can take place when traders refrain from submitting their true quantities to the LOB market, fearing a price impact. As the transaction price is frozen during a size discovery session, traders are incentivized to submit their true quantities without fearing a price impact. However, evidence on the overall effect of size discovery on market outcomes is mixed. Empirical studies argue that the introduction of size discovery sessions incentivizes traders to reduce their positions posted to the continuous LOB market, harming its liquidity. This, in turn, harms price discovery which requires up-to-date market prices. Theoretically, the evidence on the effect of workups is scarce. Antill and Duffie (2020) show that only pre-trading workups are effective. I extend this branch of research by proposing an optimal liquidation model which captures the effects of workups, market orders, and limit orders on the trader's decision-making process. My results show that traders have a high incentive to reduce the quantities sent to the LOB market and trade them in workups, instead. This effect can be considered welfare-harming, as it impedes price discovery. In simulation studies, I find that, on the other hand, the reallocation speed increases when workups are introduced. Faster reallocation can be considered welfare-enhancing, as holding unwanted inventory is costly. I extend my study with an empirical application that further demonstrates that workups can increase the transacted volume at the expanse of slightly worse average prices for the less elastic part of the market. My results should be taken into account by traders designing optimal liquidation strategies, exchanges considering alternative trading mechanisms, and regulators interested in improving market outcomes.

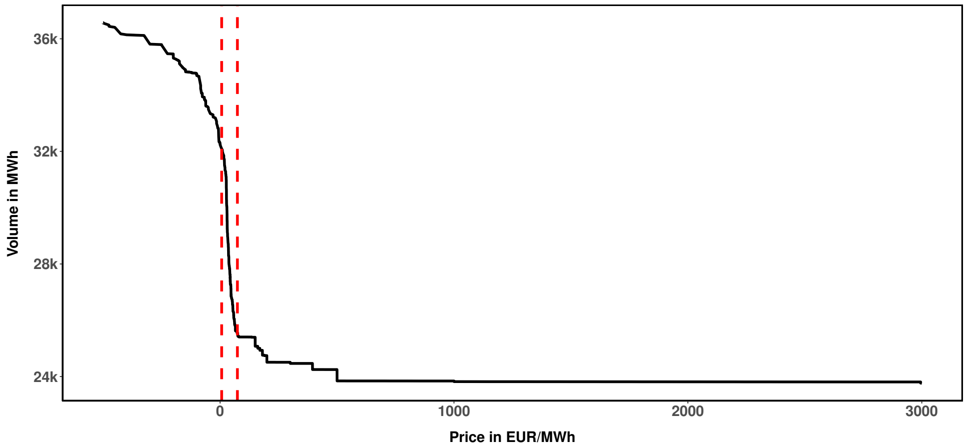
Apart from micro-structural adjustments, the efficiency of financial markets is also determined by appropriate regulations that balance the interests of market participants. In the fourth essay, I focus on disclosure as one such measure. I consider the disclosure of trades conducted by company insiders trading companies' stocks. In most financial markets, such trades have to be disclosed either in advance or, more commonly, after the transaction took place. This regulation promotes equal access to information, as company insiders might be in possession of privileged knowledge of the company. Disclosure also fosters fairness, as insiders who disclose their trades cannot, in theory, attain excess returns by utilizing privileged information. However, multiple studies show that insiders achieve excess returns in practice (Bajo and Petracci, 2006, Goncharov et al., 2013, Hodgson et al., 2020). In my essay, I concentrate on the German, the UK, and the US-American financial markets and investigate the common features of excess returns attained by insiders and their imitators, relating these to the security regulation goals. I find that in all three markets, insiders achieve excess returns and that these are correlated with insider's position within the company, the period between trade and its publication, and the trade's volume. In addition, I identify post-trade disclosure and lack of enforcement as the main causes of the ineffectiveness of current regulations. My findings equip regulators with insights on how the current disclosure rules can be improved to foster overall welfare.

Appendix

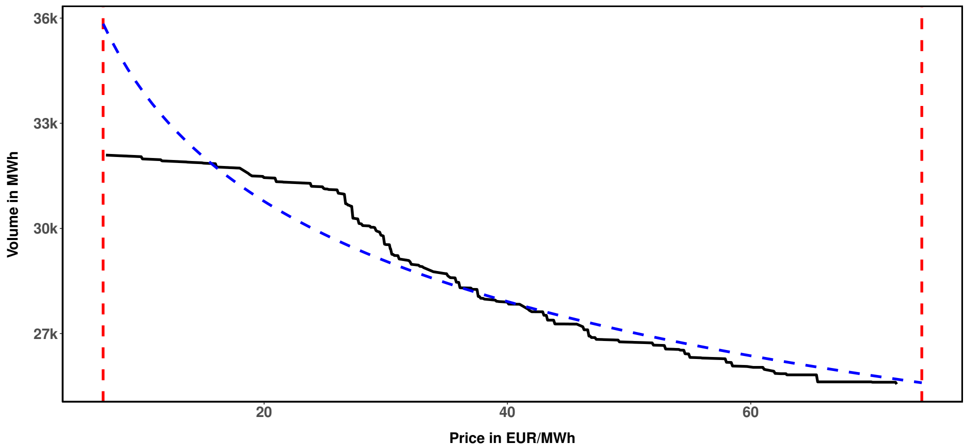
A.1 Demand Elasticity Estimation

A.1.1 Submitted Curve Sampling

As mentioned in Section 2.4, market participants are permitted to submit bids in the range from -500 EUR/MWh to 3,000 EUR/MWh. The functional form of demand changes drastically across this range of minimum and maximum bid caps. During our sample period, prices remain below 220 EUR/MWh. We apply the log-log specification in Equation (2.5) only to the part of the demand curve between the minimum and the maximum observed equilibrium price for the corresponding month. Hence if an hourly demand curve falls in the week, say, 28, we only estimate the elasticity for this curve for all observed equilibrium prices $p^* \in [\max(p^*), \min(p^*)]$, where max and min refer to the minimum and maximum prices observed in July. Figure A.1 depicts a selected demand curve for the entire price range (upper graph) and the corresponding relevant part of the demand that we seek to estimate (lower graph). The demand curve relates to hour 10 on July 11, 2019. The maximum equilibrium price in July was observed at $p^*_{max} = 74.06$ EUR/MWh, while the minimum price was observed at $p^*_{min} = 26.72$ EUR/MWh.



(a) Full aggregated curve of submitted bids.



(b) Relevant sector of the curve (seen in part (a) as dashed red lines).

Figure A.1: Quantity (y-axis) over price(x-axis) of a complete demand curve for a representative hour, along with a zoomed-in version of the relevant part of the curve. Hour 10 of July 11th, 2019, is chosen as the representative hour. Note that the whole price range from -500 EUR/MWh to 3,000 EUR/MWh is observed in the bids.

A.2 Optimal Liquidation in Continuous Markets with Workups

A.2.1 Choice of the Limit Order Fill Probability

This choice of linear/triangular distribution is primarily made to ensure analytical tractability. A quadratic (or higher-degree) (in δ) polynomial distribution implies a cubic (or higher-degree) (in Z) polynomial in the δ-controlled part of the HJB Equation 3.10. This would require a higher degree ansatz for Z and complicate the solution significantly. An exponential distribution is possible, but would require an approximation of the exp function, e.g. via $exp(-x) \approx 1 - x$ to find the optimal control δ. A uniform distribution like $\delta \sim U(0, \bar{\delta})$ would also yield a tractable solution but is considered less realistic than the assumed linear one.

A.2.2 Solving the HJB Equation

In this section, we provide a formal background for the choice of the HJB for the jump diffusion process as found in both specifications of the theoretical market model. We follow the verification Theorem 5.1 from Øksendal and Sulem (2019) and show that all conditions of the verification theorem are fulfilled in our model.

First, we embed our problem into the framework established in Øksendal and Sulem (2019), following the notation of the theorem. We set $k = \ell = 4$, $m = 1$, and $p = 3$. We set $\mathcal{S} = \mathbb{R}^3 \times (0, \infty)$ to capture that Z is positive. Further, we set $\mathcal{T} = [0, T]$, and collect the four state variables in a vector, $Y^{(u)}(t) = [C_t, P_t, \mu_t, Z_t]^\top$. Similarly, we organize the drift terms, $b(Y(t), u(t)) = [(P_t - \kappa m_t)m_t, b(\mu_t - m_t), -\kappa_f \mu_t, -m_t]^\top$, *diffusion* volatility terms, $\sigma(Y(t), u(t)) = [0, \sigma, 0, 0]^\top$, compensated Poisson jump measure terms, $\bar{N}(\mathrm{d}t, \mathrm{d}\zeta) = m(\mathrm{d}\zeta) \odot [\mathrm{d}L_t^+, \mathrm{d}L_t^-, \mathrm{d}N_t^w, \mathrm{d}N_t^\delta]^\top$ and Wiener process, $B(t) = W_t$. Further, we set the control space, $U = \mathbb{R}^3$, the objective functional, $J^{(u)}(y) = J^u(t, C, P, \mu, Z)$, stopping time, $\tau_{\mathcal{S}} = \tau$, (inventory) holding penalty, $f(Y(t), u(t)) = -\frac{\eta_f}{\tau} \int_t^\tau (Z_s)^2 \mathrm{d}s$, (cash-stock) portfolio value, $g(Y(\tau_{\mathcal{S}})) = C_\tau + Z_\tau(P_\tau - \alpha Z_\tau)$, value function, $\Phi(y) = J(t, C, P, \mu, Z)$, and *jump* volatility:

$$\gamma(Y(t^-), u(t^-), \zeta) = \begin{bmatrix} 0 & 0 & P_t w_t & P_t + \delta_t \\ 0 & 0 & 0 & 0 \\ \eta & -\eta & 0 & 0 \\ 0 & 0 & -w_t & -1 \end{bmatrix}$$

In addition to embedding our problem into the framework in Øksendal and Sulem (2019), some of the conditions of the Theorem 5.1 require an explicit formulation of the value function. Recalling all ansätze assumed in the foregoing development, the explicit value function is given by:

$$J(t, C, P, \mu, Z) = C + ZP + m_0(t) + \mu m_1(t) + \mu^2 m_2(t) + Z[\ell_0(t) + \mu \ell_1(t)] + Z^2 j_2(t), \quad \text{(A.1)}$$

subject to the terminal conditions, $0 = m_0(T) = m_1(T) = m_2(T)$.

Noting that $\mathcal{L}^\mu j_0 = 2(\eta^2\lambda_f - \mu^2\kappa_f)m_2 - \kappa_f\mu m_1$, by the same procedure as in Appendix A.2.9, we find $m_{1,2}$ as given by:

$$\begin{aligned} m_1(t) =& \frac{1}{4}\int_t^T e^{-\ \kappa_f(s-t)} \left(2\left[\frac{1}{\kappa} - \frac{\lambda_d}{j_2(s)} + \lambda_f\frac{\kappa_\ell^2}{2}\right]\ell_0(s)\ell_1(s) - \lambda_f\kappa_\ell[2+\kappa_\ell j_2(s)]\ell_1(s)\right)\mathrm{d}s, \\ m_2(t) =& \frac{1}{4}\int_t^T e^{-2\kappa_f(s-t)} \quad \left[\frac{1}{k} - \frac{\lambda_d}{j_2(s)} + \lambda_f\frac{\kappa_\ell^2}{2}\right] \quad \ell_1^2(s)\mathrm{d}s. \end{aligned} \tag{A.2}$$

And further, m_0 is given by:

$$\begin{aligned} m_0(t) = \frac{1}{4}\int_t^T \Bigg(& 8\eta^2\lambda_f m_2(s) + \left[\frac{1}{\kappa} - \frac{\lambda_d}{j_2(s)} + \lambda_f\frac{\kappa_\ell^2}{2}\right]\ell_0^2(s) \\ & + \lambda_f\kappa_\ell[2+\kappa_\ell j_2(s)]\left[\frac{1}{\kappa_\ell} - \ell_0(s) + \frac{j_2(s)}{2}\right]\Bigg)\mathrm{d}s. \end{aligned} \tag{A.3}$$

In the following, we use these implicit definitions of m_0, m_1, m_2 to prove the conditions of Theorem 5.1 of Øksendal and Sulem (2019), as an explicit calculation requires a restriction of the model parameters and is not necessary.

With these preparations, we now proceed to verify the conditions of Theorem 5.1 from Øksendal and Sulem (2019).

First, note that the optimal controls (3.23)-(3.25) are determined to satisfy the HJB (3.21), and constrained to lie in the admissible set of strategies, $\mathcal{A}$, thereby fulfilling condition (v) of the Theorem 5.1 of Øksendal and Sulem (2019). The remaining conditions are addressed as follows:

(i) Note that the image under the Lévy generator, A^u, of the value function, $\phi(y) \equiv J(t, C_t, P_t, \mu_t, Z_t)$, is given by the right-hand side of the HJB Equation (3.21) without the term $-\eta_f Z^2 \equiv$

$f(y, u)$. Using (A.1), we get:[52]

$$\overbrace{m_0' + \mu m_1' + \mu^2 m_2' + Z(\ell_0' + \mu\ell_1') + Z^2 j_2'}^{\partial_t h} + \overbrace{2(\eta^2\lambda_f - \mu^2\kappa_f)m_2 - \kappa_f\mu m_1 - Z\kappa_f\mu\ell_1}^{\mathcal{L}^\mu j} + b\mu Z$$
$$+ \sup_m\{-\kappa m^2 - m(bZ + \ell_0 + \mu\ell_1 + 2Zj_2)\}$$
$$+ \lambda_f \sup_\delta\{\quad \kappa_\ell(1 - {}^{\kappa_\ell\delta}/_2)(\delta + j_2 - \ell_0 - \mu\ell_1 - 2Zj_2)\}$$
$$+ \lambda_d \sup_w\{\quad w^2 j_2 - w(\ell_0 + \mu\ell_1 + 2Zj_2)\}, \tag{A.4}$$

where primes ($'$) denote time-derivatives. By condition (v) the optimal controls (3.23)-(3.25), $u^* \in \mathcal{A}$, satisfy the HJB Equation (3.21), $(\forall y \in \mathcal{S})\ A^{u^*}\phi(y) + f(y, u^*) = 0$, and each control, $\{m, \delta, w\}$, only indirectly affects terms outside its optimization problem via the state variables, μ and Z.[53] From this, we follow that if any other controls, $u^* \neq v \in U$, are taken, the (supremum) optimization objectives may only decrease, whilst the rest of the HJB Equation (3.21), remains unchanged: Symbolically,

$$(\forall v \in U, y \in \mathcal{S})\ A^v\phi(y) + f(y, v) \leq A^{u^*}\phi(y) + f(y, u^*) = 0$$

(ii) In case the stopping time, $\tau = T \wedge \inf\{t : Z_t = 0\} = T$, it is clear that $y \notin \mathcal{S}$ implies that $Z_t \leq 0$ and so $t = T = \tau$. But then the value function (A.1) and terminal conditions, $0 = \ell_0(T) = \ell_1(T) = m_0(T) = m_1(T) = m_2(T)$ and $j_2(T, \mu) = -\alpha$, yield

$$\phi(y) \equiv J(\tau, C_\tau, P_\tau, \mu_\tau, Z_\tau) = C_\tau + Z_\tau(P_\tau - \alpha Z_\tau) \equiv g(y)$$

In the case $\tau < T$, none of the equations, $0 = \ell_0(\tau) = \ell_1(\tau) = m_0(\tau) = m_1(\tau) = m_2(\tau) = \alpha + j_2(\tau, \mu)$ hold generally. Because $(\forall 0 < t < T)\ \tau = t$ is possible given $Z_0 > 0$, imposing these equations would require that all functions concerned (including $\alpha + j_2(t)$) vanish identically, yielding $(\forall y \in \mathcal{S})\ \phi(y) = g(y)$. In practice, once the stopping time, $t = \tau$, is reached, liquidation stops and so does time itself, i.e., whatever occurs for $t \geq \tau$ is 'irrelevant,' and in this sense, the values of mappings $f \in \{\ell_0, \ell_1, m_0, m_1, m_2, \alpha + j_2\}$ may be set arbitrarily for such times; in particular, the assignments $f \leftarrow \mathbb{1}_{[t<\tau]}f$ may be made,

[52] None of the subsequent Equations (A.16)-(A.18), (A.19), (A.22) may be used, as they were all derived by assuming at least one optimal control, whereas condition (i) must be fulfilled for all controls.
[53] That is, each control only *explicitly* enters the HJB Equation 3.21 within its own optimization problem.

so that $0 = \ell_0(\tau) = \ell_1(\tau) = m_0(\tau) = m_1(\tau) = m_2(\tau) = \alpha + j_2(\tau,\mu)$. From this, condition (ii) follows.

(iv) [And (vi).] Verification of the uniform integrability[54] of $(\forall u \in \mathcal{A}, y \in \mathcal{S})$ $\{\phi^-(Y(t))\}_{t\leq\tau}$ $[\{\phi(Y^{(\hat{u})}(t))\}_{t\leq\tau}]$ is straightforward, taking into account the following bounds:[55]

- Applying the triangle inequality to both numerator and denominator of the rightmost factor in formula (A.21) for χ,

$$|\chi| \leq \sqrt{\frac{c}{\xi}\frac{\zeta e^{2\gamma T}+1}{\zeta \qquad -1}} \equiv \bar{\chi}$$

- Recalling that $j_2 = \chi - \frac{\beta}{\xi}$ and applying the preceding and triangle inequalities,[56]

$$0 < \underline{j} \equiv \frac{\beta}{\xi} - \bar{\chi} \leq |j_2| \leq \frac{\beta}{\xi} + \bar{\chi} \equiv \bar{j}$$

- Applying these and the triangle inequalities to formulas (A.23) for ℓ_0 and ℓ_1,

$$|\ell_0| \leq \Lambda_0 \equiv T\lambda_f\kappa_\ell^2\bar{h}\left[\frac{1}{\kappa_\ell}+\frac{\bar{j}}{2}\right]$$
$$0 \leq \ \ell_1 \ \leq \Lambda_1 \equiv Tb$$

- Applying the preceding and triangle inequalities to formulas (A.2) for m_1 and m_2,

$$|m_1| \leq K_1 \equiv \frac{T}{4}\left(2\left[\frac{1}{\kappa}+\frac{\lambda_d}{\underline{j}}+\lambda_f\frac{\kappa_\ell^2}{2}\right]\Lambda_0\Lambda_1 + \lambda_f\kappa_\ell[2+\kappa_\ell\bar{j}]\Lambda_1\right)$$
$$|m_2| \leq K_2 \equiv \frac{T}{4}\qquad\left[\frac{1}{\kappa}+\frac{\lambda_d}{\underline{j}}+\lambda_f\frac{\kappa_\ell^2}{2}\right]\qquad\Lambda_1^2$$

- Similarly applying the preceding and triangle inequalities to formula (A.3) for m_0,

$$|m_0| \leq K_0 \equiv \frac{T}{4}\left(8\eta^2\lambda_f K_2 + \left[\frac{1}{\kappa}+\frac{\lambda_d}{\underline{j}}+\lambda_f\frac{\kappa_\ell^2}{2}\right]\Lambda_0^2 + \lambda_f\kappa_\ell[2+\kappa_\ell\bar{j}]\left[\frac{1}{\kappa_\ell}+\Lambda_0+\frac{\bar{j}}{2}\right]\right)$$

[54] For more details on this verification, we refer to Chapter 13 of Williams (1991), Section 4 of Chapter 5 of Gut (2013), Appendix A.4 of Bass (2011), and Section 4.5. of Bogachev (2007).

[55] It is assumed that $\alpha\xi > \beta + \gamma > 0 < c$, in order that $\zeta > 1$ and $j_2 < 0 > \chi$.

[56] In order to uniformly bound j_2 below zero, it is further assumed that $|\beta| > \xi\bar{\chi} > \gamma$, which, together with $\beta + \gamma > 0$, implies that $\beta > \gamma > 0$.

Denoting $\bar{J} \equiv \max\{K_0, K_1, K_2, \Lambda_0, \Lambda_1, \bar{j}\}$, the value function (A.1) satisfies[57]

$$|J(t,C,P,\mu,Z)| \leq C + ZP + \bar{J}\left[(1+|\mu|)(1+Z) + \mu^2 + Z^2\right] \tag{A.5}$$

The integrability of this polynomial upper bound in C, P, μ, and Z, of $(\forall u \in \mathcal{A}, y \in \mathcal{S})$ $\{\phi^-(Y(t))\}_{t\leq\tau}$ $[\{\phi(Y^{(\hat{u})}(t))\}_{t\leq\tau}]$, establishes the uniform integrability of the latter. From this, (iv) and (vi) follow.[58]

(iii) The integrability of $|\phi(Y(t))|$, i.e., that $(\forall u \in \mathcal{A}, 0 \leq t \leq T)$ $E^y|\phi(Y(t))| < \infty$, likewise follows immediately from bound (A.5). Establishing the same for $\int_0^\tau |A^u\phi(Y(s))|\mathrm{d}s$ is straightforward, taking into account the following additional bounds:[59]

- We note that $j_2' = \chi' = -\frac{4c\zeta e^{2\gamma(T-t)}}{[1-\zeta e^{2\gamma(T-t)}]^2}$,

$$|j_2'| \leq \frac{4c\zeta e^{2\gamma T}}{(\zeta-1)^2} \equiv \bar{j}'$$

- Time-differentiating formulas (A.23) for ℓ_0 and ℓ_1,

$$|\ell_0'| \leq \Lambda_0' \equiv \Lambda_0\left[\frac{1}{T} + \xi\bar{\chi} + \frac{\lambda_d}{2}\right]$$
$$|\ell_1'| \leq \Lambda_1' \equiv \Lambda_1\left[\frac{1}{T} + \xi\bar{\chi} + \frac{\lambda_d}{2} + \kappa_f\right]$$

- Similarly differentiating formulas (A.2)-(A.3) for m_0, m_1 and m_2,

$$|m_0'| \leq K_0' \equiv K_0 \quad \frac{1}{T}$$
$$|m_1'| \leq K_1' \equiv K_1\left[\frac{1}{T} + \quad \kappa_f\right]$$
$$|m_2'| \leq K_2' \equiv K_2\left[\frac{1}{T} + 2\kappa_f\right]$$

[57]The processes P, Z, and C are almost surely nonnegative.
[58]cf. the two sentences immediately following definition and formula 4.5.1 of Bogachev (2007).
[59]The parametric assumptions of footnotes 55 and 56 are retained.

- Applying the preceding and triangle inequalities to formula (A.4) for $A^u\phi(y)$,[60]

$$\frac{1}{\tau}\int_0^\tau |A^u\phi(Y(s))|\mathrm{d}s \leq K_0' + |\mu|K_1' + \mu^2 K_2' + Z(\Lambda_0' + |\mu|\Lambda_1') + Z^2\bar{j}'$$

$$+ 2(\eta^2\lambda_f + \mu^2\kappa_f)K_2 + \kappa_f|\mu|K_1 + Z|\mu|(\kappa_f\Lambda_1 + b)$$

$$+ \kappa(m^*)^2 + |m^*|(bZ + \Lambda_0 + |\mu|\Lambda_1 + 2Z\bar{j})$$

$$+ \lambda_d[(w^*)^2\bar{j} + |w^*|(\Lambda_0 + |\mu|\Lambda_1 + 2Z\bar{j})]$$

$$+ \lambda_f[\kappa_\ell(1 + {\kappa_\ell\delta^*}/2)(\delta^* + \bar{j} + \Lambda_0 + |\mu|\Lambda_1 + 2Z\bar{j})]$$

The integrability of this polynomial upper bound in C, P, μ, and Z, establishes that $(\forall u \in \mathcal{A})\ E^y \int_0^\tau |A^u\phi(Y(s))|\mathrm{d}s < \infty$ and so, altogether, condition (iii):

$$(\forall u \in \mathcal{A}, 0 \leq t \leq T)\ E^y\left[|\phi(Y(t))| + \int_0^\tau |A^u\phi(Y(s))|\mathrm{d}s\right] < \infty$$

Finally, note that Theorem 5.1 of Øksendal and Sulem (2019) requires that the value function, $\phi(y) \equiv J(t, C_t, P_t, \mu_t, Z_t) \in C^2(\mathcal{S}) \cap C(\bar{\mathcal{S}})$, where the closure is given by: $\bar{\mathcal{S}} \equiv \{Z \geq 0\} \equiv \mathbb{R}^3 \times [0, \infty)$. This is the case for the polynomial value function (A.1) of the state variables, y, i.e., the dependence on time, t, through the mappings, m_0, m_1, m_2, ℓ_0, ℓ_1, and j_2, is irrelevant for this consideration.

A.2.3 Separability of sup-terms

We want to show that:

$$\sup_{(x,y)\in\mathcal{A}_x\times\mathcal{A}_y} \{X(t,x) + Y(t,y)\} = \sup_{x\in\mathcal{A}_x} X(t,x) + \sup_{y\in\mathcal{A}_y} Y(t,y). \tag{A.6}$$

[60] Here, m^*, δ^*, and w^* respectively denote the optimal controls (3.23)-(3.25), that achieve the supremum values in their respective optimization problems, and to which the triangle inequality is then applied.

Let $x^* \in \mathcal{A}_x$ $[y^* \in \mathcal{A}_y]$ solve $\sup_{x\in\mathcal{A}_x} X(t,x)$ $[\sup_{y\in\mathcal{A}_y} Y(t,y)]$; then $(x^*, y^*) \in \mathcal{A}_x \times \mathcal{A}_y$ and

$$\sup_{x\in\mathcal{A}_x} X(t,x) + \sup_{y\in\mathcal{A}_y} Y(t,y) = X(t,x^*) + Y(t,y^*) \leq \sup_{(x,y)\in\mathcal{A}_x\times\mathcal{A}_y} \{X(t,x) + Y(t,y)\} \tag{A.7}$$

Let $(x^*, y^*) \in \mathcal{A}_x \times \mathcal{A}_y$ solve $\sup_{(x,y)\in\mathcal{A}_x\times\mathcal{A}_y}\{X(t,x) + Y(t,y)\}$; then $x^* \in \mathcal{A}_x$, $y^* \in \mathcal{A}_y$ and

$$\sup_{(x,y)\in\mathcal{A}_x\times\mathcal{A}_y} \{X(t,x) + Y(t,y)\} = X(t,x^*) + Y(t,y^*) \leq \sup_{x\in\mathcal{A}_x} X(t,x) + \sup_{y\in\mathcal{A}_y} Y(t,y) \tag{A.8}$$

Taken together, inequalities (A.7) and (A.8) yield Equation (A.6).

A.2.4 Base Model Optimal Controls

First, we derive m^*. We have that $\partial_z j = j_1 + 2Zj_2$. Substituting this into $sup_m\{-\kappa m^2 - m\partial_z j\}$ and taking the partial derivative with respect to m, we get:

$$-2\kappa m - j_1 - 2Zj_2 = 0.$$

solving for m yields:

$$m_t^* = \frac{-j_1 - 2Zj_2}{2\kappa}.$$

Similarly, to find δ^*, we substitute $j(t,Z)$ into the $\sup_\delta$ expression in (3.11) and derive with respect to δ:

$$\frac{a(-2\delta + j_1 - aj_2 + 2P_{max} + 2j_2 Z)}{P_{max}^2} = 0,$$

which yields

$$\delta_t^* = \frac{1}{2}\left(j_1 - aj_2 + 2j_2 Z + 2P_{max}\right).$$

To find w^*, we find the derivative of the $\sup_w$ term of the Equation (3.11):

$$-j_1 - 2Zj_2 + 2j_2 w = 0,$$

and solve for w:

$$w_t^* = \frac{j_1 + 2zj_2}{2j_2}.$$

A.2.5 System of PIDEs

Resubstitute optimal controls $u = (m^*, l^*, w^*)$ into (3.11) yields:

$$\begin{aligned} 0 = \partial_t j - \eta z^2 & \\ & + \frac{(j_1 + 2j_2 Z)^2}{4\kappa} \\ & + \lambda_\delta \frac{a(j_1 - aj_2 - P_{max} + 2j_2 Z)^2}{4P_{max}^2} \\ & + \lambda_w \frac{-(j_1 + 2j_2 Z)^2}{4j_2}. \end{aligned} \tag{A.9}$$

Using $\partial_t j = \partial_t j_0 + \partial_t j_1 Z + \partial_t j_2 Z^2$, we collect all terms for Z^0 first:

$$\partial_t j_0 + \frac{j_1^2}{4\kappa} + \lambda_\delta \frac{a(-j_1 + aj_2 + P_{max})^2}{4P_{max}^2} + \lambda_w \frac{-j_1^2}{4j_2} = 0. \tag{A.10}$$

Similarly, we collect Z^1:

$$\partial_t j_1 + \frac{j_1 j_2}{\kappa} + \lambda_\delta \frac{-aj_2(-j_1 + aj_2 + P_{max})}{P_{max}^2} + \lambda_w(-j_1) = 0, \tag{A.11}$$

and Z^2:

$$\partial_t j_2 - \eta + \frac{j_2^2}{\kappa} + \lambda_\delta \frac{aj_2^2}{P_{max}^2} + \lambda_w(-j_2) = 0. \tag{A.12}$$

This yields the system of PIDEs.

A.2.6 Solution to the System of PIDEs

First, note that $j_0 = 0$, because (A.10) only depends on $\partial_t j_0$ and the remainder is just a function of j_1, j_2. Next, we note that (A.11) is linear in j_1 and therefore $j_1 = 0$, as $j_1 = 0$. Last, to derive j_2, note that it is a Riccati-type PDE with constant coefficients.

Rearranging the equation yields:

$$\partial_t j_2 = \eta + \lambda_w j_2 - \left(\frac{1}{k} + \lambda_\delta \frac{a}{P_{max}^2}\right) j_2^2.$$

We set $\theta = \left(\frac{1}{k} + \lambda_\delta \frac{a}{P_{max}^2}\right)$, then $\partial_t j_2 = -\theta j_2^2 + \lambda_w j_2 + \eta$ and by solving the quadratic equation and rearranging with linear factors, we get $\partial_t j_2 = -\theta(j_2 - C_1)(j_2 - C_2)$, where

$$C_{1/2} = \frac{\lambda_w \pm \sqrt{\lambda_w^2 + 4\theta\eta}}{2\theta}.$$

Next, to avoid quadratic terms in j_2 (which would turn cubic when integrating), we decompose this equation into fractions and integrate:

$$\int \left(\frac{1}{j_2 - C_1} - \frac{1}{j_2 - C_2}\right) d\tau = \int -\theta(C_1 - C_2) d\tau,$$

to obtain j_2, and applying $j_2 = -\alpha$:

$$1 + \frac{C_1 - C_2}{j_2 - C_1} = \exp\left\{-\theta(C_1 - C_2)(T - t)\right\} \frac{\alpha + C_2}{\alpha + C_1}.$$

Therefore, we find:

$$j_2 = \frac{C_1 - C_2}{\exp\left\{-\theta(C_1 - C_2)(T - t)\right\} \frac{\alpha + C_2}{\alpha + C_1} - 1} + C_1. \tag{A.13}$$

This solution exists, because $C_{1/2}$ always exist $(\theta, \eta, \lambda_{\delta,w} > 0)$.

A.2.7 Rate of Market Orders

The rate is given by:

$$m_t^* = \frac{-Z}{\kappa} \frac{C_2 - C_1 \Lambda \exp(-\Theta(T-t))}{1 - \Lambda \exp(-\Theta(T-t))}.$$

To see the sign of the rate m^*, we first note that $0 \in [C_2, C_1]$, because $\lambda_w < \sqrt{\lambda_w^2 + 4\eta\theta}$. Wh therefore can assume $C_1 > 0$ and $C_2 < 0$. Now consider two cases:

Case 1: $\alpha > |C_2|$. In this case we know that $\Lambda > 0$ and because $C_1 > C_2$, $\Lambda \leq 1$. With $exp(-\Theta(T-t)) > 0$ by definition, the nominator of m^* is positive, as $-Z$ multiplied with a negative factor is positive. With a positive denominator the total rate is positive in this case. The trader is therefore selling throughout the entire trading session.

Case 2 $\alpha < |C_2|$: In this case, we know that $C_1 \frac{\alpha + C_2}{\alpha + C_1} < 0$. We set $C_2 - C_1 \Lambda \exp(-\Theta(T-t)) = 0$. Solving for t, we find $t = \log\left(\frac{C_2\alpha + C_1 C_2}{C_1\alpha + C_1 C_2}\right)\frac{1}{\Theta} + T$ and therefore, the root is larger than T, i.e. no change of sign takes place in $[O, T]$. Plugging in $t = T$, we find that the rate m^* is positive, and because it does not change the sign, the seller is also selling throughout the whole session.

A.2.8 Rate of Limit Orders

We know that $C_1 > C_2$ and $\theta > 0$ by construction, therefore $\Theta > 0$. We know that $C_{1/2}$ are symmetrical by construction (either around 0 or around a positive number). We also know that $\Lambda < 1$. From this, it follows that $\exp(-\Theta(T-t)) \leq 1$ and therefore the denominator of (3.17) is positive. Similarly, if $C_2 < 0$, we know that the nominator is negative. If $C_2 > 0$, using that $C_1 \Lambda \exp(-\Theta(T-t)) < C_2$, we also know that the nominator is negative.

A.2.9 Derivation of Order Flow Strategies

First, we assume κ, λ_f, η_f, κ_f, b, λ_d, κ_ℓ, η, and α to be (non)strictly positive. From their forms (A.20), we have that $a > 0$, $c \geq 0$, such that $\gamma = \sqrt{ac} \in \mathbb{R}$.

Using the ansatz, $J(t, C, P, \mu, Z) \equiv C + ZP + j(t, \mu, Z)$, j satisfies:

$$\begin{aligned} 0 = \partial_t j + \mathcal{L}^\mu j + b\mu Z - \eta_f Z^2 + \quad & \sup_m \{-\kappa m^2 - m(bZ + \partial_Z j)\} \\ & + \lambda_f \sup_\delta \{ \quad \kappa_\ell (1 - \kappa_\ell \delta/2)[\delta + j(t, \mu, Z-1) - j]\} \\ & + \lambda_w \sup_w \{ \quad j(t, \mu, Z - w) - j\}, \end{aligned} \tag{A.14}$$

with the terminal condition $j(T,\mu,Z) = -\alpha Z^2$. The optimal feedback controls, m^* and δ^*, respectively follow from the corresponding supremum first-order conditions (FOCs):

$$\begin{aligned} m^* &= \frac{(bZ+\partial_z j)}{2\kappa} \\ \delta^* &= \frac{1}{\kappa_\ell} + \frac{j - j(t,\mu,Z-1)}{2} \\ w^* &= \frac{j_1 + Zj_2}{2j_2} \end{aligned} \tag{A.15}$$

From the ansatz, $j(t,\mu,Z) \equiv j_0(t,\mu) + Zj_1(t,\mu) + Z^2 j_2(t,\mu)$, a coupled system of partial integro-differential equations follows:

$$0 = (\partial_t + \mathcal{L}^\mu)j_0 + \frac{1}{4\kappa}j_1^2 - \frac{\lambda_d}{4j_2}j_1^2 + \lambda_f \frac{\kappa_\ell^2}{2}\left[\frac{1}{\kappa_\ell} - \frac{j_1 - j_2}{2}\right]^2 \tag{A.16}$$

$$0 = (\partial_t + \mathcal{L}^\mu)j_1 + \frac{1}{2\kappa}j_1(b+2j_2) + b\mu - \lambda_d\, j_1 - \lambda_f\, \kappa_\ell^2\left[\frac{1}{\kappa_\ell} - \frac{j_1 - j_2}{2}\right] j_2 \tag{A.17}$$

$$0 = (\partial_t + \mathcal{L}^\mu)j_2 + \frac{1}{4\kappa}(b+2j_2)^2 - \eta_f - \lambda_d\, j_2 + \lambda_f \frac{\kappa_\ell^2}{2}\, j_2^2, \tag{A.18}$$

with terminal conditions, $j_1(T,\mu) = 0$ and $j_2(T,\mu) = -\alpha$.

The terms of the Equation (A.18) and the terminal condition $j_2(T,\mu) = -\alpha$, are independent of μ, thereby j_2 is independent and satisfies a Riccati Equation that may be solved exactly:

$$0 = \partial_t j_2 + \frac{1}{4k}(b+2j_2)^2 - \eta_f - \lambda_d j_2 + \lambda_f \frac{\kappa_\ell^2}{2} j_2^2 \tag{A.19}$$

Define the following parameters:

$$\begin{aligned} \xi &\equiv \frac{1}{\kappa} + \frac{\lambda_f \kappa_\ell^2}{2} \\ \beta &\equiv \frac{b}{2\kappa} - \frac{\lambda_d}{2} \\ -c &\equiv \frac{b^2}{4\kappa} - \eta_f - \frac{\beta^2}{\xi} \end{aligned} \tag{A.20}$$

The Riccati Equation (A.19) for $\chi(t) \equiv j_2 + \frac{\beta}{\xi}$ is then:

$$1 = \frac{\partial_t \chi}{c - \xi\chi^2} = \left[\frac{1}{\sqrt{c} + \sqrt{\xi}\chi} + \frac{1}{\sqrt{c} - \sqrt{\xi}\chi}\right]\frac{\partial_t \chi}{2\sqrt{c}}.$$

Integrating from t to T, and denoting $\gamma \equiv \sqrt{\xi c}$ and $\zeta \equiv \frac{\alpha-\frac{\beta}{\xi}+\sqrt{\frac{c}{\xi}}}{\alpha-\frac{\beta}{\xi}-\sqrt{\frac{c}{\xi}}}$,

$$2\gamma(T-t) = \log \frac{\sqrt{\frac{c}{\xi}}+\frac{\beta}{\xi}-\alpha}{\sqrt{\frac{c}{\xi}}-\frac{\beta}{\xi}+\alpha} - \log \frac{\sqrt{\frac{c}{\xi}}+\chi}{\sqrt{\frac{c}{\xi}}-\chi} \iff \chi(t) = \sqrt{\frac{c}{\xi}}\frac{1+\zeta e^{2\gamma(T-t)}}{1-\zeta e^{2\gamma(T-t)}} \tag{A.21}$$

We make an ansatz, $j_1(t,\mu) \equiv \ell_0(t) + \mu\ell_1(t)$[61], and subject to the terminal conditions, $\ell_0(T) = 0 = \ell_1(T)$, $\mathcal{L}^\mu j_1 = -\kappa_f \mu \ell_1$, Equation (A.17) for j_1 reads

$$0 =\partial_t \ell_0 + \left(\xi\chi - \frac{\lambda_d}{2}\right)\ell_0 - \lambda_f\kappa_\ell^2\left[\frac{1}{\kappa_\ell}+\frac{j_2}{2}\right]j_2 + \mu\left[\partial_t\ell_1 + \left(\xi\chi-\frac{\lambda_d}{2}-\kappa_f\right)\ell_1 + b\right] \tag{A.22}$$

Non-trivial solutions for the "linear" terms (not) multiplied by μ may be written:

$$\begin{aligned}\ell_0(t) =& -\lambda_f\kappa_\ell^2 \int_t^T e^{-\frac{\lambda_d}{2}(s-t)} \exp\left[\int_t^s \xi\chi(u)\mathrm{d}u\right]\left[\frac{1}{\kappa_\ell}+\frac{j_2(s)}{2}\right]j_2(s)\mathrm{d}s,\\ \ell_1(t) =& \quad b\int_t^T e^{-\left(\frac{\lambda_d}{2}+\kappa_f\right)(s-t)}\exp\left[\int_t^s \xi\chi(u)\mathrm{d}u\right]\mathrm{d}s.\end{aligned} \tag{A.23}$$

And rewriting (A.23):

$$\begin{aligned}\ell_0(t) =& -\lambda_f\kappa_\ell^2\int_t^T e^{-\frac{\lambda_d}{2}(s-t)}\frac{\zeta e^{\gamma(T-s)}-e^{-\gamma(T-s)}}{\zeta e^{\gamma(T-t)}-e^{-\gamma(T-t)}}\left[\frac{1}{\kappa_\ell}+\frac{j_2(s)}{2}\right]j_2(s)\mathrm{d}s\\ =& -\lambda_f\kappa_\ell^2\bar{\ell}_0(t,T;\lambda_d).\\ 0 \le \ell_1(t) =& \quad b\ \bar{\ell}_1(T-t) \equiv b\bar{\ell}_1(\tau')\\ \equiv& \quad \frac{b}{\zeta e^{\gamma\tau'}-e^{-\gamma\tau'}}\left[\zeta e^{\gamma\tau'}\frac{1-e^{-\left(\frac{\lambda_d}{2}+\kappa_f+\gamma\right)\tau'}}{\frac{\lambda_d}{2}+\kappa_f+\gamma} - e^{-\gamma\tau'}\frac{1-e^{-\left(\frac{\lambda_d}{2}+\kappa_f-\gamma\right)\tau'}}{\frac{\lambda_d}{2}+\kappa_f-\gamma}\right]\end{aligned} \tag{A.24}$$

One special case in which explicit solutions are available is $(\exists n \in \mathbb{N})\ \lambda_d = n\gamma$: our investigations reveal that the resulting expressions are of a "binomial" form, with increasingly many similar summands as n grows. The cases with odd n, including $\lambda_d = \gamma$ are more complicated in that considerations of complex roots arise: it is expected that the cases with even and odd naturals, or all naturals together, may satisfy some sort of recursion relation(s), e.g., integrating by parts,

[61] We propose an affine ansatz (in μ) for j_1 analogously to the quadratic one (in Z) for j: the corresponding HJB Equation (A.17), in addition to the terminal condition, $j_1(T,\mu) = 0$, is affine in μ.

but we consider the case $\lambda_d = 2\gamma$ for concreteness and simplicity:

$$\bar{\ell}_0(t,T;2\gamma) = \frac{1}{\gamma\xi^2[\zeta e^{2\gamma(T-t)}-1]}\left(\gamma^2 \log\frac{\zeta e^{2\gamma T} - \ e^{2\gamma t}}{\zeta \quad\ -1}\right. \tag{A.25}$$

$$+\frac{\zeta}{2}\left[(\beta+\gamma)\left(\frac{\beta+\gamma}{2}-\frac{\xi}{\kappa_\ell}\right)e^{2\gamma(T-t)}-\beta\left(\frac{\beta}{2}-\frac{\xi}{\kappa_\ell}\right)\right]-2t\gamma^3 \tag{A.26}$$

$$\left.-\frac{\gamma}{2}\left[(T-t)\left(\left[\gamma+\frac{\xi}{\kappa_\ell}-\beta\right]^2-\frac{\xi^2}{\kappa_\ell^2}\right)+\frac{\zeta}{2}\gamma-\zeta\left(\frac{\xi}{\kappa_\ell}-\beta\right)\right]\right) \tag{A.27}$$

Note also that the constraint $\lambda_d = n\gamma$ must be solved in conjunction with the nonlinear system (A.20):

$$\beta_{\pm n} \equiv \frac{b}{2\kappa} - \frac{n}{2}\sqrt{\left(\frac{1}{\kappa}+\frac{\lambda_f\kappa_\ell^2}{2}\right)\left(\eta_f+\frac{\beta_{\pm n}^2}{\xi}-\frac{b^2}{4\kappa}\right)}$$

$$= \frac{1}{4-n^2}\left[\frac{2b}{k} \pm n\sqrt{\left(\frac{\xi n^2}{2}-\lambda_f\kappa_\ell^2\right)\left(\frac{b^2}{2\kappa}-2\eta_f\right)+\frac{4\eta_f}{\kappa}}\right]$$

In the case considered with $\lambda_d = 2\gamma$, the limit is taken:

$$\lim_{n\to 2}\beta_{\pm n} = \pm\left[\frac{b}{4\kappa}+\frac{\xi\kappa}{2b}\left(\frac{b^2}{2\kappa}-2\eta_f\right)\right]$$

Substituting formulas (A.24) into the ansatz, $j_1(t,\mu) \equiv \ell_0(t) + \mu\ell_1(t)$, together with the result of substituting formulas (A.20) and (A.21) into the equation, $j_2 = \chi(t) - \frac{\beta}{\xi}$, both j_1 and j_2 are explicitly determined and thereby the optimal controls.

A.2.10 Slopes of an Order Book

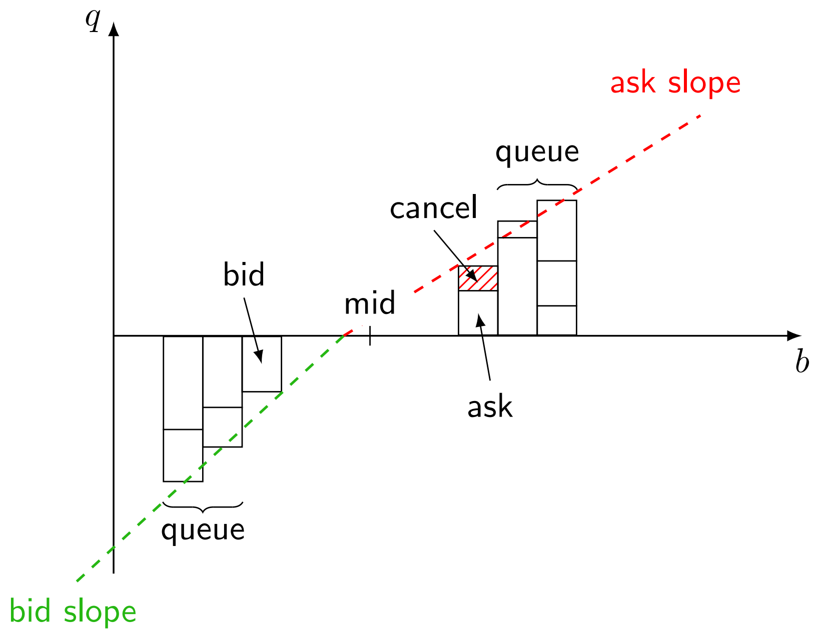


Figure A.2: visualization of the bid-ask slopes of the order book. The parameter $\psi^{bid/ask}$ is estimated from the current state of the order book by finding the average slope over cumulative orders on both sides. This graph also shows that we observe the complete queue on both sides, and order cancellations.

A.3 Fairness vs. Welfare: Disclosure in Financial Markets

A.3.1 Statistical Tests

T-test for CAARs The first employed test for the null hypothesis "$\widehat{CAAR} = \mathbb{E}[CAAR]$" is student's t-test. This test is often employed in event studies (see Dutta, 2014) and the corresponding test statistic is defined as

$$T_{t-test} = \frac{AR_{j,t}}{\sigma_{AR}}, \tag{A.28}$$

with σ_{AR} being the standard deviation of all calculated ARs. This test can be adapted to $CAAR$ (see Campbell et al., 2012):

$$T_{t-test} = \frac{CAAR[t_1, t_2]}{\sigma_{CAAR}}. \tag{A.29}$$

We apply the Kolmogorov-Smirnov and the Shapiro-Wilk tests (see Royston, 1995, Sheskin, 2020) to verify that the samples of $CAAR$s are normally distributed. Both tests reject this hypothesis at the 0.1% significance level. This applies to all trades within the scope of this work. Thus, we also apply a non-parametric test to account for the irregular distribution of $CAAR$s.

GRANK test GRANK test as introduced by Kolari and Pynnonen (2011) is a nonparametric rank test. One of the established non-parametric tests for event studies (testing abnormal returns) was introduced by Corrado (1989). It later turned out to be suited only for well-behaved situations, such as short (or even one-day) event windows and a "regular" distribution of returns (i.e., no excess kurtosis or outliers). It is furthermore very sensitive to event window length and does not account for cross-correlation of returns or event-induced volatility. GRANK test addresses these issues and is, therefore, the test of choice for our work.

Following Kolari and Pynnonen (2011) and Campbell et al. (2012), let $L_1 = T_1 - T_0 + 1$ be the estimation window length with T_0 being the first and T_1 being the last day of estimation. Further, let $L_2 = T_2 - T_1$ be the event window length with T_2 being the last day of the event window. Further, let N be the number of observations as defined above. We define the standardised abnormal return as

$$SAR_{i,t} = \frac{AR_{i,t}}{\sigma_{AR_{i,t}}}, \tag{A.30}$$

where

$$\sigma_{AR_{i,t}} = \frac{1}{L_1 - 2} \sum_{t=T_0}^{T_1} (AR_{i,t})^2. \tag{A.31}$$

Furthermore, the standardized cumulative abnormal return is

$$SCAR_{i,\tau} = \frac{CAR_{i,\tau}}{\sigma_{CAR_{i,\tau}}}. \tag{A.32}$$

Methods for computing $\sigma_{CAR_{i,\tau}}$ can be found in Campbell et al. (2012). This estimate needs to be further standardized to account for event-induced volatility:

$$SCAR^*_{i,\tau} = \frac{SCAR_{i,\tau}}{\sigma_{SCAR_\tau}}, \tag{A.33}$$

with

$$\sigma_{SCAR_\tau} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{N} (SCAR_{i,\tau} - \overline{SCAR_\tau})^2}, \tag{A.34}$$

where $\overline{SCAR_\tau} = \frac{1}{N} \sum_{i=1}^{N} SCAR_{i,\tau}$. Next, we define the generalized, standardized abnormal return

$$GSAR_{i,t} = \begin{cases} SCAR^*_i & \text{for } t_1 + 1 \leq t \leq t_1 + \tau \\ SAR_{i,t} & \text{for } \; t = T_0 + 1, \ldots, t_1, t_1 + \tau + 1, \ldots, T_2 \end{cases} \tag{A.35}$$

where $T_1 \leq t_1 \leq T_2 - \tau$ and $1 \leq \tau \leq L_2$. Next, we compute ranks for $GSAR_{i,t}$:

$$U_{i,t} = Rank(GSAR_{i,t})/(T+1) - 0.5. \tag{A.36}$$

Apart from that, all of the event period returns are combined into one. This makes the test less sensitive to varying event window lengths. Finally, the t-statistic is defined as:

$$t_{GRANK} = Z\sqrt{\frac{T-2}{T-1-Z^2}} \sim t_{T-2}, \tag{A.37}$$

where

$$Z = \frac{\overline{U_0}}{S_{\overline{U}}} \text{ with: } \overline{U_0} = \frac{1}{N}\sum_{i=1}^{N} U_{i,0} \text{ and } S_{\overline{U}} = \sqrt{\frac{1}{T}\sum_{t\in T}\overline{U_t}^2}. \tag{A.38}$$

T-test for significance of timing estimator (TE) The most straightforward way to test for significance of the timing estimator is to assume it is t-distributed. In this case,

$$T_{TE-t-test} = \frac{TE[t]}{\sigma_{TE[t]}}, \tag{A.39}$$

where $\sigma_{TE[t]}$ is the sample standard deviation. This assumption is partially fulfilled since by building averages ($CAARs$), the distribution is converging towards normal.

Bootstrapping test for significance of timing estimator (TE) Because the theoretical distribution of the TE cannot be derived in closed form, a bootstrapping strategy is employed, similar to Lyon et al. (1999). By applying this strategy, we are able to approximate the true distribution of the timing estimator and thus make a statement about its significance. In particular, we first approximate the distribution of $AR_{i,t}$ using bootstrapping. This distribution is then used to construct a distribution of $CAAR$. Then, for a given t, we approximate the distribution of $TE[t]$ by building the difference between randomly sampled $CAAR$s. In our study we set $t = 5$. To approximate $TE[5]$, we randomly and independently sample $CAAR[1, 5]$ and $CAAR[-4, 0]$. After that, we build $TE[5] = CAAR[1, 5] - CAAR[-4, 0]$. Lastly, the empirical quantiles of $TE[5]$ are used to test for significance on the corresponding level.

A.3.2 Return Metrics

In the following table, we exemplary show the results for different return metrics used in event studies to quantify the abnormal returns. The metrics are defined as follows.

Buy-and-hold abnormal return (BHAR) (Lyon et al., 1999) is defined as the difference between the buy-and-hold abnormal return of security i and the benchmark portfolio m:

$$BHAR\,[t_1, t_2] = \prod_{t=t_1}^{t_2} (1 + R_{i,t}) - \prod_{t=t_1}^{t_2} (1 + R_{m,t}). \tag{A.40}$$

Cumulative abnormal returns CAR without compounding (i.e. sum instead of product) are defined as follows:

$$CAR\,[t_1, t_2] = \sum_{t=t_1}^{t_2} AR_{i,t}. \tag{A.41}$$

Lastly, the compounded version of CAR, the CAR^* is defined as:

$$CAR^*\,[t_1, t_2] = \prod_{t=t_1}^{t_2} (1 + AR_{i,t}) - 1. \tag{A.42}$$

BHAR is considered superior to aggregated excess returns in reflecting an investor's position. However, this method of compounding fails to assess an event's impact in longer-term studies, as compounding effects drown out the event's abnormal return (Borusyak and Jaravel, 2017)

Table A.1: Comparison of return metrics

t	$BHAR$	CAR w/o comp.	CAR w/ comp.
1	2.00%	2.00%	2.00%
2	4.08%	4.00%	4.04%
3	4.12%	4.00%	4.04%
4	4.16%	4.00%	4.04%
5	4.20%	4.00%	4.04%
10	4.42%	4.00%	4.04%
20	4.88%	4.00%	4.04%
250	13.20%	4.00%	4.04%

Notes: For the calculation of return metrics we assume constant returns of the benchmark portfolio of 1% for each time interval t. The security's return is 3% in $t = 1$, 3% in $t = 2$ and follows the benchmark return of 1% thereafter for all $t > 2$.

A.3.3 Binary Regression Results

Table A.2: Binary regression results for buy trades on publication day

	Independent variable		
	Publ. Period (a)	Trade Volume (b)	Insider Level (c)
Binary regression coefficients Germany			
Publ. Period	0.008*		
	(0.004)		
Volume		−0.002	
		(0.004)	
Insider Level			0.011 ***
			(0.004)
Constant	−0.004	0.004	−0.007 *
	(0.004)	(0.003)	(0.004)
Observations	1,067	1,067	1,067
Residual Std. Error (df = 1065)	0.056	0.056	0.056
F Statistic (df = 1; 1065)	3.235*	0.352	7.350 ***
Threshold trade volume: 8,666 EUR			
Binary regression coefficients UK			
Publ. Period	−0.008***		
	(0.002)		
Volume		0.006***	
		(0.002)	
Insider Level			0.003
			(0.002)
Constant	0.013***	0.003	0.007 ***
	(0.002)	(0.002)	(0.001)
Observations	3,457	3,457	3,457
Residual Std. Error (df = 3455)	0.062	0.062	0.062
F Statistic (df = 1; 3455)	12.975***	6.791***	1.972 *
Threshold trade volume: 2,180 EUR			
Binary regression coefficients USA			
Publ. Period	−0.022***		
	(0.005)		
Volume		0.013**	
		(0.004)	
Insider Level			0.021 ***
			(0.004)
Constant	0.026***	−0.001*	−0.003
	(0.004)	(0.004)	(0.002)
Observations	844	844	844
Residual Std. Error (df = 842)	0.057	0.057	0.057
F Statistic (df = 1; 842)	22.29 ***	8.645**	24.636 ***
Threshold trade volume: 6,500 EUR			

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table A.3: Binary regression results for buy trades on transaction day

	Independent variable		
	Publ. Period (a)	Trade Volume (b)	Insider Level (c)
Binary regression coefficients Germany			
Publ. Period	0.010**		
	(0.005)		
Volume		−0.010**	
		(0.006)	
Insider Level			0.012 ***
			(0.004)
Constant	−0.004	0.013**	−0.005
	(0.004)	(0.006)	(0.004)
Observations	1,068	1,068	1,068
Residual Std. Error (df = 1066)	0.060	0.0060	0.060
F Statistic (df = 1; 1066)	4.615**	2.824*	7.255 ***
Threshold trade volume: 8,666 EUR			
Binary regression coefficients UK			
Publ. Period	−0.007***		
	(0.002)		
Volume		0.007**	
		(0.003)	
Insider Level			0.004 *
			(0.0032)
Constant	0.013***	0.003	0.007 ***
	(0.002)	(0.003)	(0.001)
Observations	3,458	3,458	3,458
Residual Std. Error (df = 3456)	0.062	0.062	0.062
F Statistic (df = 1; 3456)	9.470***	3.943**	3.797 *
Threshold trade volume: 2,180 EUR			
Binary regression coefficients USA			
Publ. Period	−0.023***		
	(0.005)		
Volume		0.010	
		(0.006)	
Insider Level			0.017 ***
			(0.004)
Constant	0.026***	−0.0002	0.003
	(0.004)	(0.006)	(0.002)
Observations	843	843	843
Residual Std. Error (df = 841)	0.055	0.055	0.055
F Statistic (df = 1; 841)	25.000***	2.316	16.833 ***
Threshold trade volume: 6,500 EUR			

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

In Table A.2, $CAAR[1,10]$ is the dependent variable in each of the binary regressions. The time period from 1 to 10 is relative to the publication day. For each of the binary variables, the state "0" is defined as: publication period < 1 day (publication period); transaction volume $< x_{0.25}$ quantile (volume); trading corporate officer is not an executive-level insider (insider level). Accordingly, the state "1" is defined as follows: publication period ≥ 1 day (publication period); transaction volume $\geq x_{0.25}$ quantile (volume); trading corporate officer is an executive-level insider (insider level). In Table A.3, the time period from 1 to 10 is relative to the transaction day.

A.3.4 Disclosure Statistics

Table A.4: DD notifications and BaFin investigations 2005-2015 (from: BaFin Annual Reports 2005-2015)

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Number of notifications	5,118	4,687	4,603	4,978	2,673	2,258	2,869	2,281	2,187	1,800	1,809
New investigations	2	11	5	7	4	3	2	7	2	2	3
Open investigations (from previous years)	152	92	24	9	9	11	10	4	8	8	7
Fines imposed	3	8	10	2	1	1	4	2	1	3	0
Open investigations (in total)	92	24	9	9	11	10	4	8	8	7	6
Highest fine [in thousand EUR]	8	5	38	16	2	4	12	35	6	15	0
Investigations closed (for legal reasons or factual reasons)	5	2	3	0	1	0	0	0	0	0	0
Investigations closed (for reasons of convenience)	54	69	7	5	0	3	4	1	1	0	4

Table A.5: DD notifications and equity trades at FWB (from: World Federation of Exchanges' monthly reports). The numbers of equity trades are rounded to millions with one decimal.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Notifications	5,118	4,687	4,603	4,978	2,673	2,258	2,869	2,281	2,187	1,800	1,809
Equity trades in [m.]	87.7	109.0	145.0	141.9	101.9	116.5	137.7	104.7	104.3	110.3	136.7

Under the pre-MAR regime, the reporting period was five business days. About half of all notifications were made within two business days, see table A.6. The median is exactly two business days. About 7 percent of trades were reported after the five-day limit. Between 2005 and 2015, 3,206 transactions were reported on average per year. In this period, only four investigations were opened per year, which resulted in three fines on average, which equates to 0.14 percent investigations and 0.1 percent fines per year. It is surprising that 7 percent of trades were reported too late, but only 0.14 percent of trades resulted in investigations, see Table A.6. These numbers allow for the conclusion that a very high number of breaches of DD rules go unpunished – even though they have been detected and found their way into the regulator's database. We refer to Table A.6 for an overview.

Table A.6: Reporting time of DD notifications in Germany

Reporting time	Reported trades	Reported trades (only trades above 5,000 EUR)
Up to and including 2 days	51%	55%
> 2 days	49%	45%
> 3 days	33%	30%
> 4 days	18%	17%
> 5 days	17%	10%
> 6 days	7%	7%
Average	4.84 days	
Median	2.00 days	
Standard deviation	22.85 days	

Notes: In this table, we use the 2iQ data as described in Table 4.1 to analyze the reporting time for German DD transactions. The sample consists of DD transactions from July 2005 until September 2015 for all DAX companies. All reported transaction types (including purchases, sells, awards, subscriptions and derivatives) are relevant for the descriptive statistics. The column "reported trades" represents all trades reported during the respective period, while column three represents only those above the de minimis threshold for reporting of DD transactions in Germany (above 5,000 EUR).

Bibliography

Abadie, Alberto and Maximilian Kasy (2019) “Choosing among regularized estimators in empirical economics: the risk of machine learning,” *Review of Economics and Statistics*, Vol. 101, pp. 743–762.

Akerlof, George A (1978) “The market for lemons: Quality uncertainty and the market mechanism,” in *Uncertainty in economics*: Elsevier, pp. 235–251.

Alfonsi, Aurélien, Antje Fruth, and Alexander Schied (2010) “Optimal execution strategies in limit order books with general shape functions,” *Quantitative finance*, Vol. 10, pp. 143–157.

Allaz, Blaise and Jean-Luc Vila (1993) “Cournot competition, forward markets and efficiency,” *Journal of Economic theory*, Vol. 59, pp. 1–16.

Alldredge, Dallin M and David C Cicero (2015) “Attentive insider trading,” *Journal of Financial Economics*, Vol. 115, pp. 84–101.

Almgren, Robert and Neil Chriss (2001) “Optimal execution of portfolio transactions,” *Journal of Risk*, Vol. 3, pp. 5–40.

Angrist, Joshua D, Kathryn Graddy, and Guido W Imbens (2000) “The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish,” *The Review of Economic Studies*, Vol. 67, pp. 499–527.

Angrist, Joshua D and Alan B Krueger (2001) “Instrumental Variables and the Search for Identification: from Supply and Demand to Natural Experiments,” *The Journal of Economic Perspectives*, Vol. 15, pp. 69–85.

Antill, Samuel and Darrell Duffie (2019) “Augmenting markets with mechanisms,”Technical report, National Bureau of Economic Research.

——— (2020) "Augmenting Markets with Mechanisms," *Review of Economic Studies*, pp. 1–55.

Arisoy, Ibrahim and Ilhan Ozturk (2014) "Estimating industrial and residential electricity demand in Turkey: a time varying parameter approach," *Energy*, Vol. 66, pp. 959–964.

Athey, Susan (2018) "The impact of machine learning on economics," in *The Economics of Artificial Intelligence: an Agenda*: University of Chicago Press.

Ausubel, Lawrence M and Peter Cramton (2010) "Using forward markets to improve electricity market design," *Utilities Policy*, Vol. 18, pp. 195–200.

Bajari, Patrick, Denis Nekipelov, Stephen P Ryan, and Miaoyu Yang (2015) "Machine learning methods for demand estimation," *American Economic Review*, Vol. 105, pp. 481–85.

Bajo, Emanuele and Barbara Petracci (2006) "Do what insiders do," *Studies in Economics and Finance*.

Baltaoglu, Sevi, Lang Tong, and Qing Zhao (2018) "Algorithmic bidding for virtual trading in electricity markets," *IEEE Transactions on Power Systems*, Vol. 34, pp. 535–543.

Bass, Richard F (2011) *Stochastic processes*, Vol. 33: Cambridge University Press.

Bechler, Kyle and Michael Ludkovski (2015) "Optimal execution with dynamic order flow imbalance," *SIAM Journal on Financial Mathematics*, Vol. 6, pp. 1123–1151.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen (2012) "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, Vol. 80, pp. 2369–2429.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014) "High-dimensional methods and inference on structural and treatment effects," *Journal of Economic Perspectives*, Vol. 28, pp. 29–50.

Beny, Laura Nyantung (2005) "Do insider trading laws matter? Some preliminary comparative evidence," *American Law and Economics Review*, Vol. 7, pp. 144–183.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995) "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, pp. 841–890.

Bhattacharya, Utpal and Hazem Daouk (2002) "The world price of insider trading," *The journal of Finance*, Vol. 57, pp. 75–108.

——— (2009) "When no law is better than a good law," *Review of Finance*, Vol. 13, pp. 577–627.

Bloomfield, Robert, Maureen O'hara, and Gideon Saar (2015) "Hidden liquidity: some new light on dark trading," *The Journal of Finance*, Vol. 70, pp. 2227–2274.

Bogachev, Vladimir I (2007) *Measure theory*, Vol. 1: Springer Science & Business Media.

Bompard, Ettore, Yuchao Ma, Roberto Napoli, and Graziano Abrate (2007) "The demand elasticity impacts on the strategic bidding behavior of the electricity producers," *IEEE Transactions on Power Systems*, Vol. 22, pp. 188–197.

Bönte, Werner, Sebastian Nielen, Niyaz Valitov, and Torben Engelmeyer (2015) "Price elasticity of demand in the EPEX spot market for electricity – new empirical evidence," *Economics Letters*, Vol. 135, pp. 5–8.

Boogen, Nina, Souvik Datta, and Massimo Filippini (2017) "Demand-side management by electric utilities in Switzerland: Analyzing its impact on residential electricity demand," *Energy Economics*, Vol. 64, pp. 402–414.

Borenstein, Severin (2005) "The long-run efficiency of real-time electricity pricing," *The Energy Journal*, Vol. 26, pp. 93–116.

Borusyak, Kirill and Xavier Jaravel (2017) "Revisiting event study designs," *Available at SSRN 2826228*.

Brandt, Michael W, Kenneth A Kavajecz, and Shane E Underwood (2007) "Price discovery in the treasury futures market," *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, Vol. 27, pp. 1021–1051.

Bris, Arturo (2005) "Do insider trading laws work?" *European Financial Management*, Vol. 11, pp. 267–312.

Budish, Eric, Peter Cramton, and John Shim (2015) "The high-frequency trading arms race: Frequent batch auctions as a market design response," *The Quarterly Journal of Economics*, Vol. 130, pp. 1547–1621.

Bushnell, James B, Erin T Mansur, and Celeste Saravia (2008) "Vertical arrangements, market structure, and competition: an analysis of restructured US electricity markets," *American Economic Review*, Vol. 98, pp. 237–266.

Buti, Sabrina, Barbara Rindi, and Ingrid M Werner (2017) "Dark pool trading strategies, market quality and welfare," *Journal of Financial Economics*, Vol. 124, pp. 244–265.

Campbell, John Y, Andrew W Lo, and A Craig MacKinlay (2012) *The econometrics of financial markets*: princeton University press.

Carrasco, Marine and Guy Tchuente (2015) "Regularized LIML for many instruments," *Journal of Econometrics*, Vol. 186, pp. 427–442.

Cartea, Alvaro and Sebastian Jaimungal (2015) "Optimal execution with limit and market orders," *Quantitative Finance*, Vol. 15, pp. 1279–1291.

——— (2016) "Incorporating order-flow into optimal execution," *Mathematics and Financial Economics*, Vol. 10, pp. 339–364.

Cenesizoglu, Tolga and Gunnar Grass (2018) "Bid-and ask-side liquidity in the NYSE limit order book," *Journal of Financial Markets*, Vol. 38, pp. 14–38.

Clastres, Cédric and Haikel Khalfallah (2015) "An analytical approach to activating demand elasticity with a demand response mechanism," *Energy economics*, Vol. 52, pp. 195–206.

Coffee Jr, John C, Hillary Sale, and M Todd Henderson (2015) "Securities regulation: cases and materials."

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe (2016) "Using big data to estimate consumer surplus: the case of uber,"Technical report, National Bureau of Economic Research.

Cont, Rama and Arseniy Kukanov (2017) "Optimal order placement in limit order markets," *Quantitative Finance*, Vol. 17, pp. 21–39.

Corrado, Charles J (1989) "A nonparametric test for abnormal security-price performance in event studies," *Journal of financial economics*, Vol. 23, pp. 385–395.

——— (2011) "Event studies: a methodology review," *Accounting & Finance*, Vol. 51, pp. 207–234.

Curato, Gianbiagio, Jim Gatheral, and Fabrizio Lillo (2017) "Optimal execution with non-linear transient market impact," *Quantitative Finance*, Vol. 17, pp. 41–54.

Dardas, Kaspar and Andre Güttler (2011) "Are directors' dealings informative? Evidence from European stock markets," *Financial Markets and Portfolio Management*, Vol. 25, pp. 111–148.

Degryse, Hans, Frank De Jong, and Vincent van Kervel (2015) "The impact of dark trading and visible fragmentation on market quality," *Review of Finance*, Vol. 19, pp. 1587–1622.

Dickgiesser, Sebastian and Christoph Kaserer (2010) "Market efficiency reloaded: why insider trades do not reveal exploitable information," *German Economic Review*, Vol. 11, pp. 302–335.

Du, Songzi and Haoxiang Zhu (2017) "What is the optimal trading frequency in financial markets?" *The Review of Economic Studies*, Vol. 84, pp. 1606–1651.

Duffie, Darrell and Haoxiang Zhu (2017) "Size discovery," *The Review of Financial Studies*, Vol. 30, pp. 1095–1150.

Dutta, Anupam (2014) "Parametric and nonparametric event study tests: a review," *International Business Research*, Vol. 7, pp. 136–142.

Dymke, Björn M (2011) *Directors' Dealings in Deutschland: Empirische Analyse der Eigengeschäfte von Unternehmensinsidern*: Springer-Verlag.

Dymke, Björn M and Andreas Walter (2008) "Insider trading in Germany. Do corporate insiders exploit inside information?" *Business Research*, Vol. 1, pp. 188–205.

Eydeland, Alexander and Krzysztof Wolyniec (2002) *Energy and power risk management: new developments in modeling, pricing, and hedging*: John Wiley & Sons.

Fabra, Natalia, David Rapson, Mar Reguant, and Jingyuan Wang (2021) "Estimating the Elasticity to Real Time Pricing: evidence from the Spanish Electricity Market," *American Economic Association Papers and Proceedings forthcoming.*

Fama, Eugene F, Lawrence Fisher, Michael C Jensen, and Richard Roll (1969) "The adjustment of stock prices to new information," *International economic review*, Vol. 10, pp. 1–21.

Farley, Ryan, Eric K Kelley, and Andy Puckett (2018) "Dark trading volume and market quality: A natural experiment," in *13th Annual Mid-Atlantic Research Conference in Finance (MARC) Paper.*

Farmer, J Doyne, Paolo Patelli, and Ilija I Zovko (2005) "The predictive power of zero intelligence in financial markets," *Proceedings of the National Academy of Sciences*, Vol. 102, pp. 2254–2259.

Federico, Salvatore, Ben Goldys, and Fausto Gozzi (2010) "HJB equations for the optimal control of differential equations with delays and state constraints, I: regularity of viscosity solutions," *SIAM Journal on Control and Optimization*, Vol. 48, pp. 4910–4937.

Fessler, Pirmin and Maximilian Kasy (2019) "How to use economic theory to improve estimators: shrinking toward theoretical restrictions," *Review of Economics and Statistics*, Vol. 101, pp. 681–698.

Fleischer, Holger (2002) "Directors' dealings," *Zeitschrift für das gesamte Handels-und Wirtschaftsrecht*, pp. 1217–1229.

Fleming, Michael J, Bruce Mizrach, and Giang Nguyen (2018) "The microstructure of a US Treasury ECN: The BrokerTec platform," *Journal of Financial Markets*, Vol. 40, pp. 2–22.

Foucault, Thierry (2010) *Limit Order Markets*: American Cancer Society, URL: `https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470061602.eqf18017`.

Foucault, Thierry, Ohad Kadan, and Eugene Kandel (2005) "Limit order book as a market for liquidity," *The review of financial studies*, Vol. 18, pp. 1171–1217.

Frei, Christoph and Nicholas Westray (2015) "Optimal execution of a VWAP order: a stochastic control approach," *Mathematical Finance*, Vol. 25, pp. 612–639.

Frey, Stefan and Patrik Sandås (2017) "The impact of iceberg orders in limit order books," *Quarterly Journal of Finance*, Vol. 7, p. 1750007.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M Shapiro (2019) "Pre-event trends in the panel event-study design," *American Economic Review*, Vol. 109, pp. 3307–3338.

Fried, Jesse M (1997) "Reducing the profitability of corporate insider trading through pretrading disclosure," *S. Cal. L. Rev.*, Vol. 71, pp. 303–392.

Friederich, Sylvain, Alan Gregory, John Matatko, and Ian Tonks (2002) "Short-run returns around the trades of corporate insiders on the London Stock Exchange," *European Financial Management*, Vol. 8, pp. 7–30.

Friederich, Sylvain and Richard Payne (2007) "Dealer liquidity in an auction market: Evidence from the London Stock Exchange," *The Economic Journal*, Vol. 117, pp. 1168–1191.

Genesove, David and Wallace P Mullin (1998) "Testing static oligopoly models: conduct and cost in the sugar industry, 1890-1914," *The RAND Journal of Economics*, pp. 355–377.

Gerbaulet, Clemens and Casimir Lorenz (2017) "dynELMOD: a dynamic investment and dispatch model for the future European electricity market,"Technical report, DIW Data Documentation.

(2019) "Machine learning in energy economics and finance: a review," *Energy Economics*, Vol. 81, pp. 709–727.

Gold, David, Johannes Lederer, and Jing Tao (2020) "Inference for high-dimensional instrumental variables regression," *Journal of Econometrics*, Vol. 217, pp. 79–111.

Goldstein, Michael A, Pavitra Kumar, and Frank C Graves (2014) "Computerized and high-frequency trading," *Financial Review*, Vol. 49, pp. 177–202.

Goncharov, Igor, Allan Hodgson, Suntharee Lhaopadchan, and Sonia Sanabria (2013) "Asymmetric trading by insiders – comparing abnormal returns and earnings prediction in Spain and Australia," *Accounting & Finance*, Vol. 53, pp. 163–184.

Gould, Martin D, Mason A Porter, Stacy Williams, Mark McDonald, Daniel J Fenn, and Sam D Howison (2013) "Limit order books," *Quantitative Finance*, Vol. 13, pp. 1709–1742.

Gregory, Alan, Rajesh Tharyan, and Ian Tonks (2009) "Stock market patterns around directors' trades: effects of director category and gender on market timing."

Guéant, Olivier, Charles-Albert Lehalle, and Joaquin Fernandez-Tapia (2012) "Optimal portfolio liquidation with limit orders," *SIAM Journal on Financial Mathematics*, Vol. 3, pp. 740–764.

Guo, Xin, Adrien de Larrard, and Zhao Ruan (2016) "Optimal placement in a limit order book," *Available at SSRN 2318220*.

Gut, Allan (2013) *Probability: a graduate course*, Vol. 75: Springer Science & Business Media.

Hajati, Maryam, Hossein Seifi, and Mohamad Kazem Sheikh-El-Eslami (2011) "Optimal retailer bidding in a DA market – a new method considering risk and demand elasticity," *Energy*, Vol. 36, pp. 1332–1339.

Hatheway, Frank, Amy Kwan, and Hui Zheng (2017) "An empirical analysis of market segmentation on US equity markets," *Journal of Financial and Quantitative Analysis*, Vol. 52, pp. 2399–2427.

Hillier, David and Andrew P Marshall (2002) "The market evaluation of information in directors' trades," *Journal of Business Finance & Accounting*, Vol. 29, pp. 77–110.

Hodgson, Allan, Michael Seamer, and Katherine Uylangco (2020) “Does stronger corporate governance constrain insider trading? Asymmetric evidence from Australia,” *Accounting & Finance*, Vol. 60, pp. 2665–2687.

Hoerl, Arthur E and Robert W Kennard (1970) “Ridge regression: biased estimation for nonorthogonal problems,” *Technometrics*, Vol. 12, pp. 55–67.

Huang, Yu Chuan and Pei Lin Tsai (2008) “Effectiveness of closing call auctions: evidence from the Taiwan Stock Exchange,” *Emerging Markets Finance and Trade*, Vol. 44, pp. 5–20.

Huberman, Gur and Werner Stanzl (2004) “Price manipulation and quasi-arbitrage,” *Econometrica*, Vol. 72, pp. 1247–1275.

Hundt, Steffen, Johanna Jahnel, and Andreas Horsch (2020) “Power Purchase Agreements and Financing Renewables: an Interdependency,” *The Journal of Structured Finance*.

Inglesi-Lotz, Roula (2011) “The evolution of price elasticity of electricity demand in South Africa: a Kalman filter application,” *Energy Policy*, Vol. 39, pp. 3690–3696.

International Organization of Securities Commissions (2010) “IOSCO: Objectives and principles of securities regulation,” URL: `https://www.iosco.org/library/pubdocs/pdf/IOSCOPD323.pdf`, Accessed on 04/05/21.

Jiang, Wei (2017) “Have instrumental variables brought us closer to the truth,” *The Review of Corporate Finance Studies*, Vol. 6, pp. 127–140.

Johann, Thomas, Tālis J Putniņš, Satchit Sagade, and Christian Westheide (2019) “Quasi-dark trading: the effects of banning dark pools in a world of many alternatives,” *Available at SSRN 3365994*.

Kaserer, Christoph and Benjamin Moldenhauer (2008) “Insider ownership and corporate performance: evidence from Germany,” *Review of Managerial Science*, Vol. 2, pp. 1–35.

Kastl, Jakub (2017) “Recent advances in empirical analysis of financial markets: industrial organization meets finance,” *Honoré, B., Pakes A., Piazzesi, M. and L. Samuelson (editors) Advances in Economics and Econometrics*, Vol. 2, pp. 231–270.

King, Jessica, Martin H Schmidt, and Richard Stehle (2015) “Do Insiders and Their Imitators Trade Profitably? Index-Specific Evidence from Germany,” *Available at SSRN 2620169).*

Knewtson, Heather S, Richard W Sias, and David A Whidbee (2010) "Style timing with insiders," *Financial Analysts Journal*, Vol. 66, pp. 46–66.

Kolari, James W and Seppo Pynnonen (2011) "Nonparametric rank tests for event studies," *Journal of Empirical Finance*, Vol. 18, pp. 953–971.

Kratz, Peter and Torsten Schöneborn (2018) "Optimal liquidation and adverse selection in dark pools," *Mathematical Finance*, Vol. 28, pp. 177–210.

Lai, Yu-Sheng (2018) "Dynamic hedging with futures: a copula-based GARCH model with high-frequency data," *Review of Derivatives Research*, Vol. 21, pp. 307–329.

Lakonishok, Josef and Inmoo Lee (2001) "Are insider trades informative?" *The Review of Financial Studies*, Vol. 14, pp. 79–111.

Lee, Inmoo, Michael Lemmon, Yan Li, and John M Sequeira (2011) "The effect of regulation on the volume, timing, and profitability of insider trading."

Lenkey, Stephen L (2014) "Advance disclosure of insider trading," *The Review of Financial Studies*, Vol. 27, pp. 2504–2537.

Lijesen, Mark G (2007) "The real-time price elasticity of electricity," *Energy economics*, Vol. 29, pp. 249–258.

Lyon, John D, Brad M Barber, and Chih-Ling Tsai (1999) "Improved methods for tests of long-run abnormal stock returns," *The Journal of Finance*, Vol. 54, pp. 165–201.

MacKinlay, A Craig (1997) "Event studies in economics and finance," *Journal of economic literature*, Vol. 35, pp. 13–39.

Manne, Henry G (1966) *Insider trading and the stock market*: Free Press.

——— (1984) "Insider trading and property rights in new information," *Cato J.*, Vol. 4, pp. 933–957.

Mansur, Erin T (2007) "Upstream competition and vertical integration in electricity markets," *The Journal of Law and Economics*, Vol. 50, pp. 125–156.

Maume, Philipp and Martin Kellner (2017) "Directors' Dealings unter der EU-Marktmissbrauchsverordnung: Placebo oder Paradigmenwechsel?" *Zeitschrift für Unternehmens-und Gesellschaftsrecht*, Vol. 46, pp. 273–311.

McLean, R David (2012) "Fooled by compounding," *The Journal of Portfolio Management*, Vol. 38, pp. 108–116.

Mittal, Hitesh (2008) "Are You Playing in a Toxic Dark Pool?: a Guide to Preventing Information Leakage," *The Journal of Trading*, Vol. 3, pp. 20–33.

Nasdaq Nordic, contract specification (2021) "Nasdaq Nordic, power contracts specification," `https://www.nasdaq.com/docs/2021/03/15/Joint-Appendix-2-Contract-Specifications.pdf`, Accessed on 05/05/21.

Nasdaq Nordic, ITCH protocol (2021) "Nasdaq Nordic, ITCH Protocol Specification," `http://www.nasdaqomx.com/digitalAssets/ 106/106127_itch_protspec_nordic_va2224.pdf`, Accessed on 05/05/21.

Neuhoff, Karsten, Nolan Ritter, Aymen Salah-Abou-El-Enien, and Philippe Vassilopoulos (2016) "Intraday markets for power: discretizing the continuous trading?" *Available at SSRN 2723902*.

Nimalendran, Mahendrarajah and Sugata Ray (2014) "Informational linkages between dark and lit trading venues," *Journal of Financial Markets*, Vol. 17, pp. 230–261.

Obizhaeva, Anna A and Jiang Wang (2013) "Optimal trading strategy and supply/demand dynamics," *Journal of Financial Markets*, Vol. 16, pp. 1–32.

Ocker, Fabian and Vincent Jaenisch (2020) "The way towards European electricity intraday auctions – status quo and future developments," *Energy Policy*, Vol. 145, p. 111731.

Øksendal, Bernt and Agnès Sulem (2019) *Applied Stochastic Control of Jump Diffusions*: Springer.

Okui, Ryo (2011) "Instrumental variable estimation in the presence of many moment conditions," *Journal of Econometrics*, Vol. 165, pp. 70–86.

Pagnotta, Emiliano S and Thomas Philippon (2018) "Competing on speed," *Econometrica*, Vol. 86, pp. 1067–1115.

Pancs, Romans (2014) "Workup," *Review of Economic Design*, Vol. 18, pp. 37–71.

Parlour, Christine A and Duane J Seppi (2008) "Limit order markets: a survey," *Handbook of financial intermediation and banking*, Vol. 5, pp. 63–95.

Putniņš, Tālis J (2013) "What do price discovery metrics really measure?" *Journal of Empirical Finance*, Vol. 23, pp. 68–83.

Rahman, Dewan, Barry Oliver, and Robert Faff (2020) "Evidence of strategic information uncertainty around opportunistic insider purchases," *Journal of Banking & Finance*, Vol. 117, p. 105821.

Riordan, Ryan and Andreas Storkenmaier (2012) "Latency, liquidity and price discovery," *Journal of Financial Markets*, Vol. 15, pp. 416–437.

Rostek, Marzena and Marek Weretka (2015) "Dynamic thin markets," *The Review of Financial Studies*, Vol. 28, pp. 2946–2992.

Roşu, Ioanid (2009) "A dynamic model of the limit order book," *The Review of Financial Studies*, Vol. 22, pp. 4601–4641.

——— (2019) "Fast and slow informed trading," *Journal of Financial Markets*, Vol. 43, pp. 1–30.

——— (2020) "Liquidity and information in limit order markets," *Journal of Financial and Quantitative Analysis*, Vol. 55, pp. 1792–1839.

Royston, Patrick (1995) "Remark AS R94: a remark on algorithm AS 181: the W-test for normality," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 44, pp. 547–551.

Sağlam, Mehmet, Ciamac C Moallemi, and Michael G Sotiropoulos (2019) "Short-term trading skill: An analysis of investor heterogeneity and execution quality," *Journal of Financial Markets*, Vol. 42, pp. 1–28.

Sannikov, Yuliy, Andrzej Skrzypacz et al. (2016) "Dynamic trading: price inertia and front-running," *Preprint*.

Seligman, J., T.L.L. Loss, and T. Paredes (2011) *Fundamentals of Securities Regulation, Sixth Edition*: WOLTERS KLUWER LAW & BUSINESS, URL: `https://books.google.de/books?id=3pz2tAEACAAJ`.

Sensoy, Ahmet (2019) "Commonality in ask-side vs. bid-side liquidity," *Finance Research Letters*, Vol. 28, pp. 198–207.

Seyhun, H Nejat (1986) "Insiders' profits, costs of trading, and market efficiency," *Journal of financial Economics*, Vol. 16, pp. 189–212.

Sheskin, David J (2020) *Handbook of parametric and nonparametric statistical procedures*: CRC Press.

Shmueli, Galit and Otto R Koppius (2011) "Predictive analytics in information systems research," *MIS quarterly*, pp. 553–572.

Siu, Chi Chung, Ivan Guo, Song-Ping Zhu, and Robert J Elliott (2019) "Optimal execution with regime-switching market resilience," *Journal of Economic Dynamics and Control*, Vol. 101, pp. 17–40.

Tebroke, Hermann-Josef and Alexander Wollin (2005) "Directors' Dealings, Informationseffizienz und Handelsstrategien," *Die Unternehmung*, pp. 31–54.

Tibshirani, Robert (1996) "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288.

van der Veen, Reinier A.C. and Rudi A. Hakvoort (2016) "The electricity balancing market: exploring the design challenge," *Utilities Policy*, Vol. 43, pp. 186–194.

Varian, Hal R (2014) "Big data: New tricks for econometrics," *Journal of Economic Perspectives*, Vol. 28, pp. 3–28.

Vayanos, Dimitri (1999) "Strategic trading and welfare in a dynamic market," *The Review of Economic Studies*, Vol. 66, pp. 219–254.

Ventoruzzo, Marco (2015) "Comparing insider trading in the United States and in the European Union: history and recent developments," *European Company and Financial Law Review*, Vol. 11, pp. 554–593.

Wang, Wei and Nanpeng Yu (2019) "A machine learning framework for algorithmic trading with virtual bids in electricity markets," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, IEEE.

Weiler, Lothar and Oliver Tollkühn (2002) "Die Neuregelung des "directors' dealing" nach dem Vierten Finanzmarktförderungsgesetz," *Der Betrieb*, Vol. 55, pp. 1923–1927.

Williams, David (1991) *Probability with martingales*: Cambridge university press.

Wilson, Robert (2002) "Architecture of power markets," *Econometrica*, Vol. 70, pp. 1299–1340.

World Federation of Exchanges (2021) "Statistics on worldwide exchanges," URL: `https://www.world-exchanges.org/our-work/statistics`, Accessed on 04/05/21.

Zhu, Song-Ping and Guiyuan Ma (2018) "An analytical solution for the HJB equation arising from the Merton problem," *International Journal of Financial Engineering*, Vol. 5, p. 1850008.

Zou, Hui (2006) "The adaptive lasso and its oracle properties," *Journal of the American statistical association*, Vol. 101, pp. 1418–1429.