Market-Maker Short Selling: A Necessary Evil?

Yerim Kim[*] Kyong Shik Eom[†] Baeho Kim[‡]

First Version: August 25, 2024
This Version: May 7, 2025[§]

Abstract

This paper investigates the impact of market-makers' short-selling activities, focusing on the single-stock futures (SSFs) market in Korea during the short-selling ban (March to April 2021), when only market makers were permitted to sell the underlying stocks short. Employing supervised and unsupervised machine-learning techniques, we classified market-makers' short-selling activities into three categories—aggressive, reluctantly compliant, and willingly compliant—and applied overlap propensity-score weighting to mitigate confounding biases. Our results reveal that SSFs market-makers' aggressive short selling significantly improved liquidity, reduced volatility, and enhanced price efficiency, which is remarkable given the restrictive regulations that limited short-selling volumes. However, these short-selling activities did not alleviate the backwardation effect in the spot markets, suggesting that the regulatory restrictions limited SSFs market makers from fully realizing the benefits of short selling. This study is the first to empirically examine the economic role of short-selling activities conducted by SSFs market makers and their influence on market quality in futures and spot markets. Our findings emphasize the need for flexible regulations that fully harness the cross-market benefits of their short selling, while minimizing the risk of market and political abuses.

Keywords: Single-stock futures market-maker, short selling, overlap propensity-score weighting, machine-learning technique, market quality

JEL Codes: G1, G2, C4, C8

[*]411 LG-POSCO Building, Korea University Business School, Seoul, 02841, South Korea, Phone: +82-2-3290-2626, Email: `yerim21c@korea.ac.kr`.

[†]530 Evans Hall #3880, University of California, CRMR at Department of Economics, Berkeley, CA 94720, USA, Phone: +1-415-271-7228, Email: `kseom2@gmail.com`.

[‡]Corresponding Author, 520 LG-POSCO Building, Korea University Business School, Seoul, 02841, South Korea, Phone: +82-2-3290-2626, Fax: +82-2-922-7220, Email: `baehokim@korea.ac.kr`.

[§]We are grateful to Robert Anderson, Kiwoong Byun, Thummim Cho, Sunwoo Hwang, Heebum Lee, Jaeyeon Lee, James Park, Dean Ryu, Sungbin Sohn, Shu-Feng Wang and participants in a seminar at Korea University Business School, the 2024 Fall Korean Finance Association Conference, and the Korea-Japan Joint Symposium of Statistics and Data Science 2024 at Sungshin Women's University; for comments that greatly improved our results. This research is supported by Korea University Business School Research Grant.

1. Introduction

In financial markets, short selling refers to selling securities that a trader does not own. It is a trading strategy used by the trader to profit when he/she expects the price of a security to fall, or to hedge against potential price volatility of related derivative securities he/she owns. To engage in short selling, a trader must first borrow securities from a broker-dealer or other institution that must be returned on a specified date. The trader then sells the shares, hoping to repurchase them at a lower price to return to the lender. Academics and industry professionals generally share the view that short selling can reduce stock price overvaluation and improve market quality while not causing the stock price to fall further (e.g., Atmaz, Basak, and Ruan, 2024; Edwards, Reed, and Saffi, 2024; Khan, 2024; for more detailed information, see Section 2). However, the narrative differs when it comes to naked short selling, which involves selling stocks without borrowing them in advance. As naked short selling exposes the trader to settlement risk, which can ultimately lead to systemic risk in financial markets, all major stock markets worldwide, including Korea, prohibit this type of short selling.

Let us look at a recent case in Korea. On October 16, 2023, the Financial Supervisory Service (FSS) reported that it had caught two foreign-owned global investment banks habitually engaging in naked short selling for an extended period. In a follow-up, the FSS announced it needed to improve short-selling rules and infrastructure to create a level playing field between foreign and institutional investors and retail investors. In preparation, it banned short selling on all listed securities from November 6, 2023. Exceptions were made for spot and futures market makers (including Liquidity Providers (LPs))[1] because market makers endogenously

[1]There are two types of market makers in the major global equity spot markets: US and European. In general, the former is required to provide liquidity in quote-driven markets, such as the NYSE (Specialist/Designated Market Makers (DMM)) and Nasdaq (dealers), where market making by one or more market makers, i.e., DMMs or dealers, is a listing requirement. The latter is required to provide liquidity in order-driven markets, such as Deutsche Börse and Euronext, where investment firms (IFs) can contract with exchanges to execute market-making strategies using algorithmic trading for all tradable instruments. Alternatively, listed firms lacking liquidity, e.g., small and mid-cap firms, can voluntarily contract with an IF, i.e., a Designated Sponsor (DS) on Deutsche Börse or an LP on Euronext, to promote liquidity (Deutsche Börse AG, 2023). Although these IFs are all called market makers (or sometimes DMMs), they differ in terms of legal status, obligations, and rights. Nonetheless, existing studies do not clearly distinguish between them. Meanwhile, the Korea Exchange (KRX) stock markets are order-driven and also have both types of market makers. However, while an LP is the same as the European type, a market maker in the KRX is distinguished from both US and European types, in that when the liquidity of a listed firm falls below a certain pre-determined level, a market maker can select that firm and make a market for it. For the purpose of this paper, we study market makers in the KRX futures and stock markets, with their unique characteristics.

provide liquidity between buyers and sellers and engage in intermediate transactions. Without market makers, transactions could become insufficient and inefficient, making it difficult for financial markets to function properly, even in order-driven markets. Nonetheless, retail investors strongly opposed this decision and vociferously demanded that short selling should also be banned for market makers. The reasoning was that if they were exempted, their short selling could still be abused as an illegal conduit and/or suppress stock prices, even with a complete ban on short selling. These voices of retail traders that the rules should be applied equally to all investors were reflected in subsequent rulemaking and infrastructure improvement plans. In fact, retail traders in the Korean stock markets have been calling for fairness with institutional investors for over a decade, often to the point of overreaching whenever a short-selling controversy arises. This anomaly has emerged because policymakers have been politically sensitive to the interests and demands of retail investors, who account for a substantial portion of the market, rather than consistently aligning financial market rules and regulations with their policy philosophy.[2]

So, is market-makers' short selling in the Korean stock markets a social evil undermining market health, as retail investors and some policymakers claim? This paper seeks to answer this question. Specifically, it empirically analyzes how active market-making by single-stock futures (SSFs) market makers through short selling in the spot markets on the Korea Exchange (KRX) affects the economic function of the SSFs market and, by extension, the stock markets. Economic function refers to whether market makers' short-selling activities contribute to market quality (including liquidity and pricing accuracy) in the SSFs market and to price efficiency in futures and spot markets. Thus, this analysis aims to shed light on the general mechanisms and importance of SSFs market-makers' short selling, which is particularly important considering that most existing research focuses on the impact of short selling on the spot or options markets.

To fulfill this paper's purpose, we must identify a period when only SSFs market makers were allowed to sell short in the stock markets, as account-level data is virtually unavailable. In this sense, the period during which the FSS banned short

[2] The November 2023 short-selling ban prevented the Korean stock markets' inclusion in the 2024 Morgan Stanley Capital International Index (MSCI) Developed Markets (World) index. For more information on the ban and subsequent developments, please see the following: Jaewon Kim, 2023, "South Korea's Stock Short-selling Ban Raises Political Questions: Experts, Media See Potential Negative Impact of 'Populism' Ahead of April Election," *Nikkei Asia*, (November 27); Ji-Won Choi, 2024, "Why Korea's So Tough on Short Selling: Following November Ban, 9 Foreign Banks Busted for Illegal Stock Short Sales Worth $155m," *The Korea Herald*, (May 19); Yeon-Woo Lee, 2024, "Korea's MSCI Inclusion Faces Setback due to Extended Short Selling Ban," *The Korea Times*, (June 16).

selling due to the COVID-19 pandemic is a truly relevant period for our analysis. To calm the sharp drop in the stock markets due to the COVID-19 pandemic, the FSS banned short selling on all listed securities from March 16, 2020, and then partially lifted the ban on May 3, 2021, to allow short selling of KOSPI 200 index and KOSDAQ 150 index constituents, while retaining the ban on other stocks. Therefore, we utilize the two months immediately before the partial lifting as the 'short-selling ban period' (March 2 to April 30, 2021) for the analysis; for more detailed discussions, see Section 3. We do this for two reasons. First, SSFs' underlying stocks are constituents of the KOSPI 200 index or the KOSDAQ 150 index. Second, only SSFs market makers could 'practically' sell short during the short-selling ban period, allowing us to independently analyze the effect of SSFs market-makers' short selling.[3]

We used Trade and Quote (TAQ) data for SSFs and their underlying stocks for the analysis. Guided by the extant literature, we constructed a set of market microstructure variables to assess market quality in the SSFs market such as liquidity, volatility, and price efficiency. Those variables include effective spread, price impact, Amihud's (2002) illiquidity ratio, realized volatility, variance ratio, execution shortfall (slippage or implementation shortfall), and pricing error variability. Additionally, as the cross-market variable, we utilized the market basis gap to capture the discrepancy between the market and theoretical basis, normalized by the fair basis, where the basis is the difference between futures and spot prices.

Overall, the short-selling activities of SSFs market makers were less than expected; however, we observed notable variations in their intensity. Hence, our empirical study employed sophisticated machine-learning techniques to analyze SSFs market-makers' short-selling activities during the short-selling ban period. The first step involves clustering trading activities based on the data density of short selling relative to the total trading volume of each stock. We utilized the Data-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, an unsupervised machine-learning technique that allows for the effective identification of patterns and clusters without prior assumptions about the data structure. The DBSCAN clustering algorithm revealed that it was optimal to group short-selling activities into three clusters, based on the thresholds of 0.44% and 0.53% of short selling relative to total

[3]During this period, short selling by index futures market makers was prohibited. Short selling by stock market makers was permitted, but its amount was almost zero for two reasons: first, most stocks in which markets were being made were illiquid (see footnote 1), so the amount that could be shorted was extremely limited. Second, stock market makers could only sell short firms that they had previously shorted, making short selling virtually impossible. Conversely, the underlying stocks of SSFs market makers were firms representing the KOSPI market (the main board) or the KOSDAQ market (the growth market), so there was little overlap with the stocks of stock market makers.

trading volume.

This finding, which identifies a narrow band between 0.44% and 0.53%, aligns with the '0.5% guideline' imposed during the short-selling ban.[4] Under this guideline, KRX's internal policy limited short selling to no more than 0.5% of daily trading volume. Although SSFs market makers would normally use short selling for various purposes such as hedging their futures positions,[5] they faced strict restrictions on short selling in the KRX stock markets. While most market makers adhered to this guideline, some exceeded it for various reasons. Indeed, SSFs market makers engaged in a broad spectrum of short-selling activities, similar to their behavior under normal conditions. Grounded by these observations, we categorized market-makers' short-selling activities into three groups: (i) the 'aggressive' group, which voluntarily exceeded the guideline; (ii) the 'reluctantly compliant' group, which marginally adhered to the guideline's limits despite their reluctance; and (iii) the 'willingly compliant' group, whose short-selling activities were not constrained by the guideline.

To examine the economic function of SSFs market-makers' short-selling activities, we designed two binary treatment group comparisons: (i) 'aggressive vs. (reluctantly plus willingly) compliant,' and (ii) 'reluctantly compliant vs. willingly compliant.' For each binary comparison, propensity scores were calculated using the Gradient Boosting Machine (GBM) classification algorithm, a supervised machine learning technique that can account for nonlinear effects. The selection of the GBM algorithm was informed by an out-of-sample comparison of classification performance with other methods, including Random Forest (RF), eXtreme Gradient Boosting (XGB), and Logistic Regression (LR). The findings demonstrated that GBM exhibited superior performance with statistical significance under five-fold cross-validation, attaining the highest out-of-sample accuracy while avoiding overfitting.

We then applied the overlap propensity-score weighting method (Thomas, Li, and

[4]The 0.5% guideline was implicit, so it is not formally documented, but can be confirmed by interviews with market participants at the time. The rationale for the 0.5% was that it was small enough that retail investors would not blame market makers for short selling. During the COVID-19 pandemic, retail investors were upset with stock price decline. Policymakers were cautious of this, so they instructed market makers to minimize short selling to allay retail investors' concerns. Guideline violations could result in indirect sanctions (e.g., a disadvantage for extending a market-maker contract).

[5]Manaster and Mann (1996) argue that "market makers in the futures market are not just passive order fillers, but profit-seeking individuals with heterogeneous levels of information and/or trading skill." According to them, market makers execute short selling not only for hedging but also for profit-seeking. See Daures-Lescourret and Monias (2023) for the varying motivations of market makers operating in cross-listed stocks, rather than stocks and futures.

Pencina, 2020) to each binary comparison using the estimated propensity scores to compute the Average Treatment Effect (ATE), mitigating confounding selection bias in analyzing the short-selling effects of SSFs market makers. To ensure robustness, we additionally conducted a panel-data analysis that addresses endogeneity concerns (e.g., Eom, Ok, and Park, 2007; Bryzgalova, Pavlova, and Sikorskaya, 2023), incorporating a binary short-selling dummy and market microstructure variables to provide further insights into the economic impact of market-makers' short-selling activities in each binary comparison.

Our empirical results show that short selling by SSFs market makers improved liquidity, reduced volatility, and enhanced price efficiency, indicated by reductions in effective spread, realized volatility, variance ratio, and execution shortfall. More specifically, we observe that 'reluctantly compliant' short selling by SSFs market makers did not lead to significant improvements in market quality. Furthermore, 'aggressive' short selling increased liquidity while reducing volatility and improving price efficiency in the futures market. These results are particularly surprising, even though total short positions were small as a result of a restrictive regulatory environment. However, an analysis of the market basis gap indicated that short selling did not alleviate backwardation effects in the spot markets, suggesting that the stringent restrictions on short selling by SSFs market makers confine their contributions to the futures market, preventing spillover into the spot markets. Overall, our findings suggest that regulatory limitations have prevented SSFs market makers from engaging in short selling to the degree sufficient to fully realize its benefits.

This paper makes important contributions, both academically and institutionally. First, from an academic perspective, this is the first paper to analyze the economic impact of SSFs market-makers short-selling activities on futures and spot markets. Despite the importance of the subject, to the best of our knowledge, there is no research on market making for SSFs in Korea or abroad, let alone studies on SSFs market-makers' short selling. Furthermore, there are no studies on the pricing efficiency of the stock market (i.e., the underlying asset market) when SSFs market makers actively utilize short selling to make markets. Next, from an institutional perspective, there is insufficient research on the role of market makers and the pros and cons of their short-selling activities in global financial markets, including in Korea. The Korean financial derivatives market is internationally important, as it is the largest in Asia after those of India and China and larger than Eurex (based on the number of contracts in 2023). Furthermore, its SSFs market ranks third globally (see Section 3.1).

Studies such as this can enhance the effective regulation of market makers.

The remainder of this paper is organized as follows. Section 2 discusses the related literature and our contributions, especially regarding market makers and short-selling constraints/bans. Section 3 briefly explains the KRX SSFs market and its market maker, and presents the data and methodology, focusing on the overlap propensity-score weighting method. Section 4 presents and interprets the empirical results and their implications. Section 5 concludes with our findings and suggestions for further research.

2. Contributions to Related Literature

This paper contributes to two areas of research in finance: market makers and short-selling constraints/bans. This section outlines each of these, along with our contributions.

2.1. Market Makers

This paper empirically analyzes issues related to market making in a futures market by examining its links to spot markets. Ours is the first academic paper to address this topic in a modern financial market environment characterized by electronic trading. Few papers have investigated the period before the futures exchange became electronic, where *locals*, i.e., professional market makers, led futures trading in the pit (e.g., Manaster and Mann, 1996; Locke and Sarkar, 2001). Although they provide interesting implications about the 'typical' characteristics of market makers in futures markets, applying their findings to today's market environments is challenging because the market structure has changed dramatically.

In the electronic trading environment, most existing literature relevant to this paper discusses the 'equity spot' market 'empirically.'[6] These studies mostly report that (Designated) Market Makers, or simply (D)MMs, whether in the US or Europe,

[6] Market makers are indispensable players in the theoretical modeling of the market microstructure. Thus, since Demsetz (1968), there has been substantial theoretical research on them from various aspects, but only two studies, Venkataraman and Waisburd (2007) and Bessembinder, Hao, and Zheng (2015), are relevant to our paper. Both modeled the role of (D)MMs (more precisely, LPs) for small and mid-cap stocks with poor liquidity in the secondary market and highlighted their positive role. Both only dealt with European (D)MMs, which leaves out US (D)MMs. Another notable research strand is liquidity provision. When liquidity is extremely depleted due to exogenously given (aggregate) liquidity or volatility shocks (e.g., the 2008 Global Financial Crisis), the increased risk to liquidity supply causes liquidity providers to withdraw the (associated) liquidity supply itself, increasing their expected returns (in other words, this was the main driver of the Global Financial Crisis). Here, liquidity providers broadly include not only (D)MMs, but also HFT firms and even individual traders (e.g., Nagel, 2012; Drechsler, Moreira, and Savov, 2020).

increase trading volume and improve market quality, as evidenced by decreases in (effective) bid-ask spread and intraday volatility and an increase in market depth, price improvement (e.g., Bessembinder, Hao, and Zheng, 2020). Meanwhile, under normal market conditions, high-frequency trading (HFT) firms, in addition to (D)MMs, have voluntarily performed a significant portion of the market-making activities since the mid-to-late 2000s. Clark-Joseph, Ye, and Zi (2017) and Coughlan and Orlov (2023) examine this phenomenon in the stock and futures markets, respectively; however, there is still a lack of research distinguishing between the roles of (traditional) mandatory (D)MMs and new voluntary HFT firms when trading is intertwined among numerous market centers.

As mentioned, this paper empirically studies market makers in the electronic trading futures market. Therefore, it is worth taking a closer look at Coughlan and Orlov (2023), who implicitly consider HFT as voluntary market making and analyze agricultural, energy and metals, and financial[7] futures markets. Using account-level data, their results show a strong positive relationship between greater HFT participation and improved market quality. However, an increase in the share of aggressive liquidity-consuming HFT, e.g., directional trading, adversely affects market quality. Nevertheless, they argue that in their sample period, the former effect outweighs the latter and that, overall, HFT improves market quality in futures markets.

In addition, some studies examine market makers in options markets (e.g., Huh, Lin, and Mello, 2015; Moussawi, Xu, and Zhou, 2024) and cryptocurrency markets (e.g., Malinova and Park, 2024). Limiting their study to options market makers (OMMs) hedging adverse-selection risk through the stock market, Huh, Lin, and Mello (2015) report that OMMs' hedging trades unintentionally increase bid-ask spreads in both markets, which happens because "the [OMMs'] hedging inadvertently acts as a conduit of private information." Moussawi, Xu, and Zhou (2024) explore OMMs' hedging and arbitrage behaviors by analyzing their market making together with an ETF and an associated index option. For this, they consider the occurrence of order flow shocks in the ETF market as the informational moment and attempt to identify how OMMs hedge and secure fleeting arbitrage opportunities in ETF markets at informational moments. Furthermore, Moussawi, Xu, and Zhou (2024) empirically find that market makers incentivize (institutional) investors by adjusting the bid-ask and volatility spreads of the options market so that these investors can act as counterparties and perform their own synthetic hedges through options

[7]Here, financials include the Euro FX, E-mini S&P 500, and 10-Year Treasury Note.

while conducting fleeting arbitrage in the ETF market. An automated market maker (AMM) is a market-making algorithm determining the price in digital asset markets. Malinova and Park (2024) show that well-designed AMMs have the potential to offer improvements over traditional financial markets.[8] However, there are no papers on market makers in electronic trading futures markets. Given the lack of research on stock market makers, research on futures market makers has been further neglected by scholars. Furthermore, there may have been difficulties in obtaining at least TAQ data on stocks and futures together, and the voices of retail investors may not have been as loud in other countries as in Korea, so even if there were related issues, they may not have attracted scholars' attention.

2.2. Short-Selling Constraints or Bans

This paper analyzes short selling by market makers in the SSFs market following a ban on short selling in the spot markets. In Korea, such research is possible as short selling is allowed for market makers in the futures market even when short selling is prohibited for everyone else, which has not been discussed in the existing literature. Therefore, of the numerous short-selling studies,[9] only those focusing on short-selling constraints/bans and their issues related to market quality or derivatives markets are relevant to this paper. Even then, most of these existing studies focus on the stock markets, and the remaining studies examine the economic impact of the options markets.[10] Unless otherwise noted, the discussion of short selling below is limited to the stock markets.

Most existing theoretical studies argue that short-selling constraints lead to the overvaluation of stock prices, which results in information inefficiency in the stock markets (e.g., Miller, 1977; Atmaz, Basak, and Ruan, 2024). This is consistent with the view that short-selling constraints limit negative information from being incorporated into prices. Conversely, several studies argue that short-selling constraints do not inherently cause stock prices to be overvalued (e.g., Jarrow, 1980; Diamond and Verrecchia, 1987). The divergent results among these theoretical studies are due

[8] Considering the coexistence of competing market-making structures, the traditional centralized exchanges (CEXs) and the decentralized exchanges (DEXs) with AMMs, Aoyagi and Ito (2024) theoretically extend the traditional spot market-maker discussion to cryptocurrency markets.

[9] For a recent comprehensive literature review of short-selling constraints/bans, see Atmaz, Basak, and Ruan (2024), Edwards, Reed, and Saffi (2024), and Khan (2024).

[10] In addition to stocks, options (rather than futures) were emphasized as related research focuses on the US financial markets; however, there are few such studies. In the US, trade occurs overwhelmingly in (individual stock) options rather than (individual stock) futures; the opposite is true in Asia and Europe (World Federation of Exchanges, 2022, 2024) (see Section 3.1 for related information).

to differences in the assumptions made in the models about perceived information uncertainty, expectations homogeneity and revisions, trading motives, changes in the demand and supply for risky assets, and so on (Khan, 2024).

The inaccuracy of data for measuring short-selling constraints makes empirical research challenging. However, as in theoretical research, empirical studies have maintained that short-selling constraints delay incorporating negative information in stock prices. Thus, most argue that the prices of associated stocks are overvalued and that market quality, including liquidity, deteriorates. Ultimately, academics generally believe that short selling contributes to information/price efficiency and liquidity. Empirical studies on short-selling constraints/bans related to this paper can be broadly categorized into the following two strands.[11]

Short-Selling Constraints/Bans and Their Relationship to Market Quality. The selling demand resulting from short selling and the buying demand due to the redemption of borrowed shares create additional liquidity. Moreover, unlike the theoretical studies that disagree about the informedness of short selling,[12] empirical studies almost uniformly report that short sellers are informed traders (e.g., Dechow et al., 2001; Boehmer, Jones, and Zhang, 2008; Boehmer et al., 2020). The increase in informed trading due to short selling can positively or negatively impact liquidity and price discovery (e.g., Blau and Whitby, 2018). First, the increase in short sellers' informed trading can improve price discovery by reducing price uncertainty, narrowing bid-ask spreads, and increasing information efficiency. Conversely, it may also increase adverse-selection risk, which could worsen price discovery by widening bid-ask spreads and reducing information efficiency. When applied to short-selling constraints/bans, which are directly relevant to this paper, the results of these numerous empirical analyses mostly support the former. That is, the literature consistently reports that short-selling constraints/bans decrease liquidity and information efficiency, thereby deteriorating market quality (e.g., Marsh

[11] The impact of short selling on stock prices and returns, or short-selling fees, is another major topic. Numerous studies show that stocks with more severe short-selling constraints are likelier to be overvalued, resulting in negative *ex-post* excess returns. This phenomenon is also robust when investors' expectations are more heterogeneous (e.g., Jones and Lamont, 2002; Asquith, Pathak, and Ritter, 2005; Diether, Lee, and Werner, 2009b; Beneish, Lee, and Nichols, 2015; Prado, Saffi, and Sturgess, 2016; Sikorskaya, 2023, among numerous others).

[12] Theoretical studies examining the impact of short selling under normal market conditions suggest that short sellers contribute to efficient prices (e.g., Miller, 1977; Diamond and Verrecchia, 1987), but prices become less informative when short sellers are predatory traders (Brunnermeier and Pedersen, 2005) or use short-selling strategies as manipulation (Goldstein and Guembel, 2008).

and Payne, 2012; Choe and Lee, 2012; Beber and Pagano, 2013).[13] Here, this deterioration in market quality means increases of spreads, volatility, and price impact, and poorer price discovery. The literature also reports that short-selling constraints/bans generally have little effect on the price level. Policymakers use the claim that short-selling causes excessive declines in a falling market to postulate that constraining/banning short selling can prevent stock price decline (Edwards, Reed, and Saffi, 2024). However, this 'price support' argument has been somewhat unconvincing in academic circles (e.g., Boehmer, Jones, and Zhang, 2013; Beber and Pagano, 2013).[14]

Several studies also focus on the Uptick Rule, which is not a short-selling ban but a measure that makes it more difficult. This rule prevents short sellers from driving stock prices to fall further or gives priority to buyers in a falling market. The results of studies that lifted this rule, even temporarily, are also somewhat informative for this paper (e.g., Diether, Lee, and Werner, 2009a; Boehmer and Wu, 2013). Overall, these studies suggest the Uptick Rule reduces short selling but does not prevent it. It may have a minimal price effect, but if any, the effect is not enough to prevent price decline (e.g., Jain, Jain, and McInish, 2012). Alternatively, it may worsen market quality due to slow price discovery (e.g., Boehmer and Wu, 2013).

Short-Selling Constraints/Bans and The Existence of Derivatives Markets. The existence of derivatives markets may (e.g., Figlewski and Webb, 1993; Danielsen and Sorescu, 2001) or may not (e.g., Battalio and Schultz, 2011) weaken the effectiveness of short-selling constraints/bans. Consider the options markets as an example. In the former case, options traders are considered informed traders (e.g., Black, 1975; Easley, O'Hara, and Srinivas, 1998; Hao, Lee, and Piqueira, 2013). When short selling is constrained/banned, these traders can be expected to view the options markets as an alternative to short selling and trade more there, which is

[13] As HFT firms voluntarily engage in market making in contemporary financial markets, the research on their short selling provides helpful insights for our paper. During periods of extreme volatility when short selling was banned and liquidity was depleted (e.g., the short-selling ban period on financial stocks in the US), short selling by HFT traders negatively impacted liquidity, while short selling by non-HFT traders positively impacted it (Brogaard, Hendershott, and Riordan, 2017). Furthermore, even under normal circumstances, there are liquidity-demanding short sellers who are likely to be informed HFT traders (Comerton-Forde, Jones, and Putnins, 2016).

[14] Conversely, Chang, Cheng, and Yu (2007) analyzed stocks added to or removed from the short-selling list on the Hong Kong Exchange (HKEX). They found a price-support phenomenon in the 'not shortable' category. Brunnermeier and Oehmke (2014) provide a theoretical rationale for why temporary short-selling constraints mostly do not produce price effects (i.e., price support), but only in exceptional cases.

even more possible given the existing empirical findings that short sellers are also informed traders (e.g., Boehmer et al., 2020). The latter case shows that the US stock markets experienced a significant decline in options trading volume on banned stocks during the 2008 short-selling ban (e.g., Battalio and Schultz, 2011; Grundy, Lim, and Verwijmeren, 2012). The ban on short selling made it impossible for put-option underwriters to take short positions on the underlying stocks to hedge their option positions. As such, the results of empirical analyses in the options markets, where most studies are concentrated, are mixed. This paper considers a situation where SSFs market makers are shorting in the spot markets when short selling is prohibited, so they naturally contribute to trading volume. Considering this, the contribution of this paper is to identify the cross-market price efficiency of the SSFs market by analyzing the basis, which is the difference between futures and spot prices.

Compared to the options markets, research on the futures markets is rare. Jiang, Shimizu, and Strong (2022) report that during the 2008 short-selling ban, both SSFs trading volume of short-selling banned stocks and the contribution of SSFs trading to underlying stock price-discovery increased significantly, mitigating the deterioration in the market quality of the banned stocks.[15] Hence, when short selling is banned, futures can improve the market quality of the stock markets. Thus, Jiang, Shimizu, and Strong's (2022) findings closely relate to our paper. However, our paper differs as we analyze SSFs market-makers' short selling and its impact on the quality of the futures market itself and the market-making ability (including short selling) through basis analysis with the spot markets. Accordingly, our paper provides insights into the general mechanisms of the SSFs market-makers' economic role: short selling in the underlying spot markets (for purposes such as hedging, risk management, or arbitrage transactions) facilitates market making and improves SSFs market quality.

As we analyze the Korean financial markets, we should highlight one more peculiar issue, in addition to the short-selling issues in global stock markets. It concerns the negative role of short selling by foreign investors. This issue has arisen because foreign investors account for an overwhelming proportion of short selling in

[15]Other studies related to SSFs focus on which markets (SSFs or spot) lead price discovery, using Hasbrouck's (1995) information share (*IS*) as a measure. Although the data is rather old (January 2003 to July 2005), in the US markets, the average *IS* of the SSFs markets was 24.4~26.1%, so price discovery in the spot markets dominated that in the SSFs markets (Shastri, Thirumalai, and Zutter, 2008). Conversely, in the National Stock Exchange of India (NSE), when new information reached the markets, the *IS* of the SSFs markets was 55%, and if the information was negative, it increased to 61% (Aggarwal and Thomas, 2019). Therefore, in these situations, price discovery occurred more in the SSFs markets than in the spot markets, based on the argument that the main driver for this was short-selling constraints in the spot markets.

the Korean stock markets, e.g., 88% in 2010 and 71% in 2023. However, Wang and Lee (2015) and Eom, Binh, and Kim (2011) report that although foreign investors in the Korean stock markets are informed traders with short-run stock price predictability, their short selling does not cause stock price declines.

3. Empirical Design and Methodology

The objective of this paper is to examine whether SSFs market-makers' short selling positively supports overall market quality or results in market dysfunction as retail investors and some politicians have argued. Addressing this question requires rigorous empirical methods that can effectively handle endogenous effects while accounting for the unique features of the market structure and data under analysis. In this section, we first briefly introduce the structure of the KRX financial derivatives markets, their international status, and the processing of our sample data. Next, we explain how we use machine learning techniques to group market-making activities by intensity, how we calculate market quality variables, and why we use the overlap propensity-score weighting method for causal inference.

3.1. Structure of the KRX Financial Derivatives Markets

The KRX financial derivatives markets ranked 10$^{\text{th}}$ in the world (2.04 billion contracts) in combined futures and options trading volume (based on the number of contracts traded) in 2023.[16] It is similar in size to the Shanghai Futures Exchange (SFE, 2.06 billion contracts, 9$^{\text{th}}$) and slightly larger than the Eurex (1.92 billion contracts, 11$^{\text{th}}$).

Due to a lack of diversification in the product portfolio, the trading volume of the KRX financial derivatives markets is heavily concentrated on KOSPI 200 futures and options, as well as SSFs. As of 2023, KOSPI 200 futures ranked 8$^{\text{th}}$ among world stock index futures, while KOSPI 200 weekly index options ranked 6$^{\text{th}}$ (Thursday expiration) and 10$^{\text{th}}$ (Monday expiration) among world stock index options (World Federation of Exchanges, 2024). The status of the KRX SSFs market in the world SSFs markets is even more surprising. As of 2023, it accounted for 17.8% of the world SSFs markets, ranking third globally after Borsa Istanbul (BIST, 36.5%) and B3 SA (28.1%). Furthermore, the

[16] Overwhelmingly, the largest was the NSE (84.82 billion contracts), followed by B3 SA (Brasil, Bolsa, Balcão, 8.31 billion contracts), the Bombay Stock Exchange (BSE, 5.87 billion contracts), and the Chicago Mercantile Exchange (CME, 5.42 billion contracts). The remaining six of the top ten markets were similar in size. Futures Industry Association (FIA), 2023, "Global Futures and Options Volume Hits Record 137 Billion Contracts in 2023."

trading volume (based on the number of contracts traded) and open interest are also steadily increasing. In particular, Samsung Electronics ranked first in the world in SSFs trading volume, and half of the top 10 SSFs traded in 2023 were Korean SSFs.[17] Therefore, the KRX SSFs market can provide vital insights for world SSFs markets.

The SSFs market makers in Korea make markets for the nearest month futures contracts for all listed SSFs. In addition, starting four trading days prior to the last trading day of the nearest month contract, they make markets for the subsequent nearest contracts. The market maker for each SSFs may be monopolistic or duopolistic, depending on the performance evaluation ranking.

Market makers have been selected by the KRX using quantitative and non-quantitative evaluations over the past two years. Quantitative evaluations are based on the degree of excess of obligations, bid-ask spreads and quantity, and trading volume, while non-quantitative evaluations are based on compliance with regulations and the market-making operation plan for the following period. Market makers who disagree with the market-making operation plan will be excluded from the selection. If the short-selling limit is violated,[18] penalty points will be imposed, a warning may be issued, or contracts may be terminated.

3.2. Causal Inference Framework: $y = f(x)$

Our empirical analysis centers on examining how market-makers' short-selling intensity influences critical market quality dimensions. Conceptually, this functional relationship can be described as the interaction between x, denoting the market-makers' short-selling intensity, and y, representing the set of market quality variables. The underlying causal mechanism, $f(\cdot)$, reflects the extent to which the market makers' activities impact various aspects of market quality, constituting the primary focus of our empirical study.

Addressing confounding bias is crucial for accurately assessing the impact of SSFs market-makers' short-selling activities. This bias occurs when an external factor, known as a *confounder*, affects both the independent and dependent variables, resulting in misleading relationships between them. For example, if the underlying stocks targeted by market makers for (possibly excessive) short selling are significantly influenced by confounders, such as industry-specific factors affecting

[17] In the US, the SSFs markets are small compared to the single-stock options (SSOs) market, which accounts for 81.7% of world SSOs trading volume. One of the main reasons for this is that the SSFs markets were formed late due to jurisdictional conflicts between the Commodity Futures Trading Commission (CFTC) and the Securities and Exchange Commission (SEC).

[18] Based on this provision, market makers can be fined if they violate the implicit 0.5% guideline.

liquidity, it could give a false impression unless the industry distribution is balanced between the treatment and control groups.[19] In the presence of selection bias, the increased liquidity may be solely due to market-makers' short-selling behavior, when in fact, it may be partially or entirely attributed to these confounding factors.

We employ the short-selling ban period as a proxy to isolate market-makers' short-selling activities and investigate their causal effect on market quality. To address the challenge of defining the intensity of these activities (x), we implement an unsupervised machine-learning clustering algorithm to categorize short-selling activities into distinct intensity groups. Market quality (y) is assessed across the three key dimensions of liquidity, volatility, and price efficiency, providing a comprehensive understanding of its relationship with short selling. To establish causality, $f(\cdot)$, the selection bias is mitigated by calculating propensity scores using a sophisticated supervised machine-learning classification algorithm, ensuring well-defined treatment groups. Recognizing the limitations of traditional regression-based methods in capturing the causal effects of short selling, we employ overlap propensity-score weighting to improve causal inference and the robustness of our results.[20]

3.3. Market-Makers' Short-Selling Intensity: x

Our dataset comprises 146 SSFs contracts, including 124 underlying stocks from the KOSPI market (the main board) and 22 underlying stocks from the KOSDAQ market (the growth market) on the KRX. This dataset spans 44 trading days during the short-selling ban period from March 2 to April 30, 2021. We use this period as it represents a time with no significant market disruptions, as Figure A.1 in the Appendix demonstrates. We analyze market-makers' short-selling activities during the short-selling ban period, leveraging daily short-selling volume data from the KRX to isolate and examine their dynamics.

To capture the extent of short selling that originated from the SSFs market, we cluster their short-selling activities based on the short-selling ratio, defined as the proportion of short-selling volume by SSFs market makers to total trading volume for each underlying stock per day. We apply the DBSCAN method to identify

[19] As supported by our empirical results, stocks in the financial sector tend to be easier targets for short selling due to their high leverage, sensitivity to economic conditions, and transparent regulatory reporting. They are also exposed to systemic risks, reputational sensitivities, and sector-specific challenges such as interest rate fluctuations, making them more volatile and attractive to market makers for short selling. These factors, combined with the sector's reliance on trust and susceptibility to negative sentiment, make stocks in the financial sector prime targets during periods of uncertainty or stress.

[20] We also present a series of panel regression results to validate our empirical findings.

these clusters, an unsupervised machine-learning algorithm designed to group data points based on their density (Ester et al., 1996). DBSCAN is particularly effective in identifying clusters of arbitrary shapes and distinguishing outliers as noise, allowing for a clearer separation between varying levels of short-selling intensity. Unlike other clustering algorithms, DBSCAN does not require the number of clusters to be predefined and performs well even with large and irregularly distributed datasets.

3.4. Market Quality Metrics: y

The variables used to evaluate market quality in the SSFs market are broadly classified into three main groups associated with liquidity, volatility, and price efficiency. The liquidity measures include *effective spread*, *price impact*, and *Amihud illiquidity ratio*, while the price efficiency measures include *variance ratio*, *execution shortfall*, and *pricing error variability*. In addition, we consider *market basis gap*. This variable reflects the price accuracy by measuring the discrepancy in market bases, defined as the difference between futures and spot market prices, from the theoretically fair basis.[21]

The primary dataset used to construct these variables is TAQ data from the KRX SSFs market, covering the short-selling ban period from March 2 to April 30, 2021. This period provides a unique context for directly examining the impact of short-selling activities by SSFs market makers. We analyze all SSFs because market makers have a de facto market-making obligation to them. We exclude transactions outside the regular trading session of the KRX SSFs market (before 9:00 a.m. and after 3:45 p.m.) from the analysis. Following existing literature, we only use the nearest month futures contracts, except for the four trading days preceding the last trading day of the nearest month. For the four excluded trading days, we use the data from the next-nearest contracts. In addition to TAQ data, we also utilize daily data from the SSFs market and the underlying stock markets, as well as daily short-selling volume and stock loan balance data provided by the KRX over the same period.

3.4.1. Liquidity measures

Effective Spread. *Effective spread* is the most important measure of liquidity. It measures the cost of executing a market order by comparing the execution price to the midpoint of the bid-ask spread (Brogaard, Hendershott, and Riordan, 2017). Its

[21] The theoretically fair basis is defined as the difference between the market basis ($F_{\text{market}} - S_{\text{market}}$) and the theory basis ($F_{\text{theory}} - S_{\text{market}}$), where F is the futures price and S is the underlying stock price.

definition is given by

$$Effective\ Spread_{i,t} = \frac{1}{K_{i,t}} \sum_{k=1}^{K_{i,t}} \frac{I_{i,k,t}(p_{i,k,t} - mid_{i,k,t})}{mid_{i,k,t}},$$

where i denotes each SSFs contract,[22] k denotes each trade, t denotes each trading day, $K_{i,t}$ represents the number of trades of the i-th contract on day t, $p_{i,k,t}$ and $mid_{i,k,t}$ represent the transaction price and the mid-price of the best bid and ask prices at the k-th trading moment, respectively, and $I_{i,k,t}$ is an indicator variable that takes the value 1 for a buy order and -1 for a sell order. The effective spread provides a measure of trading costs that capture the 'hidden' costs incurred by traders. A higher effective spread indicates less liquidity in the market, implying that traders face higher costs when executing trades. This may be due to wider bid-ask spreads or the market's inability to absorb large orders without significant price impacts.

Price Impact. The *price impact* variable captures the extent to which a specific transaction affects a stock price. It is calculated as the difference between the effective spread and the realized spread to measure the gross losses of liquidity demanders to better informed traders due to adverse selection (Hendershott, Jones, and Menkveld, 2011). Specifically, we calculate this variable as

$$Price\ Impact_{i,t} = \underset{k}{\text{Average}} \left(I_{i,k,t} \cdot \frac{\widehat{mid}_{i,k,t} - mid_{i,k,t}}{mid_{i,k,t}} \right),$$

where $\widehat{mid}_{i,k,t}$ denotes the mid-price of the best bid and ask prices of the i-th contract realized five minutes after the k-th trade on day t. The price impact increases when there is a large transaction in the market, especially with low liquidity and/or low depth.

Amihud Illiquidity Ratio. *Amihud illiquidity ratio* (*Amihud*) measures market liquidity by evaluating how much a stock's price changes relative to trading volume. Conceptually, it is "a proxy for the price impact" (Brauneis et al., 2021). Adopting

[22] A 'contract' refers to an SSFs contract with the same underlying asset, regardless of its expiration month. Since market makers in Korea provide liquidity for the nearest month SSFs contracts and begin making markets for the next-nearest contracts four trading days before expiration, they may trade contracts with different maturities at the same time. To reflect this, contracts with the same underlying asset but different expiration months are treated as a single contract unit in the analysis.

Amihud's (2002) methodology, we calculate *Amihud illiquidity ratio* for each SSFs as

$$Amihud(h)_{i,t} = \frac{1}{T_h} \sum_{j=1}^{T_h} \frac{|r_{i,j,t}(h)|}{V_{i,h,t}},$$

where i denotes each contract, t denotes each trading day, T_h is the number of the h-hour intervals during a day, $r_{i,j,t}(h)$ is the j-th return over the h-hour interval, and $V_{i,h,t}$ denotes the trading volume during the h-hour interval. A higher ratio signifies reduced liquidity, indicating that the market has less ability to absorb transactions without substantial price movements.[23] In our analysis, *Amihud illiquidity ratio* was calculated using 15-minute intervals.[24]

3.4.2. Volatility measures

Realized Volatility. *Realized volatility* (*RV*) quantifies the variability of stock returns over a specified intraday interval, providing a comprehensive measure of price fluctuations and market activity (e.g., Andersen et al., 2003). Our definition of *RV* for contract i during the h-hour interval on day t is given by

$$RV(h)_{i,t} = \sum_{j=1}^{T_h} r^2_{i,j,t}(h),$$

where the h-hour interval was set to one, five, and ten minutes in our analysis.

3.4.3. Price efficiency measures

Variance Ratio. *Variance ratio* (*VR*), proposed by Lo and MacKinlay (1988), is a metric used to measure the accuracy of price discovery in financial markets by examining intraday returns over various time intervals to detect mean reversion or momentum in stock prices. Our *VR* calculation for each contract i on each trading day t follows the equation given by

$$VR(h)_{i,t} = \frac{\overline{\sigma}^2_{i,t}(h)}{h \times \overline{\sigma}^2_{i,t}(1)},$$

[23] Following Kang and Jeong (2018), we multiplied the ratio by 10^9 to account for the trading volume unit.

[24] We also calculated *Amihud illiquidity ratio* over various time intervals including one, five, ten, thirty, and forty-five minutes, and found that the empirical results are not statistically sensitive to the selection of these intervals. Additional details are available upon request.

where $\overline{\sigma}^2_{i,t}(h)$ represents the estimated variance for the h-hour return. $VR\,(h)_{i,t}$ symbolizes the ratio of the long-horizon (h-hour) return variance to h times the short-horizon ($\overline{\sigma}^2_{i,t}(1)$; 1-hour) return variance. When $VR\,(h)_{i,t}$ is greater (less) than 1, it suggests that market prices underreact (overreact) to information within the 1-hour short horizon, resulting in increased (decreased) long-term return volatility. In our analysis, we use the absolute value of $1 - VR\,(h)_{i,t}$, where a larger value implies decreased price efficiency. We calculate *VR* with $h = 4$ (Eom, Seon, and Chang, 2010).

Execution Shortfall. *Execution shortfall* (*ES*) measures the difference between the benchmark price and the final execution price of a trade (Haslag and Ringgenberg, 2023). Our *ES* formula is given by

$$ES_{i,t} = \frac{1}{K_{i,t}} \sum_{k=1}^{K_{i,t}} \frac{VWAP_{i,k,t} - P^0_{i,t}}{P^0_{i,t}} \times I_{i,k,t} \ ,$$

where $VWAP_{i,k,t}$ (Volume-Weighted Average Price) denotes the average price at which contract i is traded on trading day t, weighted by the volume of each trade k, $K_{i,t}$ is the number of trades of the i-th contract on day t, and $P^0_{i,t}$ is the benchmark price, defined as the opening price of the i-th contract on day t. The indicator variable $I_{i,k,t}$ takes the value 1 for a buy order and -1 for a sell order. A larger execution shortfall indicates that the trade is executed less efficiently, resulting in a greater deviation from the intended price, reflecting that the market conditions or trading strategies may have led to suboptimal trade execution.

Pricing Error Variability. *Pricing error* (*PE*) measures the temporary deviation between the actual transaction price and the efficient price, reflecting market frictions unrelated to information. *PE* is calculated using intraday transaction or quote data and captures temporary deviations from a random walk. Motivated by Hasbrouck (1993), we decompose the observed log transaction price, $p_{i,k,t}$, into an efficient price component, $m_{i,k,t}$, which follows a random walk, and a stationary pricing error component $s_{i,k,t}$ (i.e., $p_{i,k,t} = m_{i,k,t} + s_{i,k,t}$, for each contract i on each trade k on day t). The efficient price is defined as the expected value of a stock conditional on all available information, including both public data and private information inferred from order flow. Hasbrouck (1993) assumes the pricing error $s_{i,k,t}$ to be a zero-mean, covariance-stationary process, which may exhibit serial correlation or correlation with the innovations in the random walk of the efficient price. Our analysis employs *pricing error variability*, defined as the daily standard deviation of the contract-specific *PE*, to quantify the extent of the deviations from the efficient price and serves as a measure of price efficiency (Boehmer and Wu, 2013). A larger pricing error variability indicates lower price accuracy and decreased market quality.

3.4.4. Cross-market quality measure

Market Basis Gap. The *market basis gap* variable quantifies the excess of the market basis, which is the difference between the futures price and the spot price, as a relative metric compared to its theoretically fair basis. The formula for each contract i on each day t is expressed as

$$\textit{Market Basis Gap}_{i,t} = \frac{1}{K_{i,t}} \sum_{k=1}^{K_{i,t}} \left(\frac{\textit{Market Basis}_{i,k,t} - \textit{Theoretical Basis}_{i,k,t}}{\textit{Theoretical Basis}_{i,k,t}} \right),$$

where i denotes each contract, k denotes each trade, and $K_{i,t}$ is the number of trades of the i-th contract on day t. A positive market basis gap indicates the excess of contango, while a negative value reflects the excess of backwardation relative to the theoretically fair basis level. In our analysis, we standardize this variable on a cross-sectional basis by subtracting its mean and scaling it to have a unit sample variance, thereby normalizing it to facilitate comparison across different scales.

3.5. Causality Identification Methodology: $f(\cdot)$

3.5.1. Conventional approaches and their limitations

A randomized controlled trial (RCT) is the gold standard for causal inference (e.g., LeLorier et al., 1997; Duflo, Glennerster, and Kremer, 2006), as randomization ensures a balanced distribution of known and unknown confounders, allowing an unbiased assessment of the ATE. However, observational studies often lack this advantage,[25] leading to imbalanced confounders and potentially spurious ATE estimates.

While multiple regression can address some confounders, it may struggle to compare groups due to the risk of overfitting when dealing with numerous covariates. Instrumental variables offer another approach to establishing causality, but finding suitable instruments can be challenging, as stock-level characteristics influencing market-makers' short selling also generally affect outcome variables. Furthermore, a key assumption in difference-in-difference (DiD) analysis is that a treatment is

[25] In finance, a few papers (e.g. Levy, 2022) have just begun attempting to use RCT in a limited specific area.

randomly assigned (Roth et al., 2023), which is unlikely to hold in our study. A standard DiD model assumes that a market maker's decision to select a particular stock for (excessive) short selling, as well as the extent of this activity, is random. However, as noted above, this assumption rarely holds in practice, making it essential to address the selection bias in how stocks are assigned to a treatment.

3.5.2. Overlap propensity-score weighting

When appropriate instrumental variables are unavailable, researchers often turn to statistical techniques such as propensity score matching and weighting. These methods require calculating the probability (i.e., propensity score) of an individual stock receiving a specific "treatment" (i.e., market-makers' short selling) based on observable covariates. As proposed by Rosenbaum and Rubin (1983), the goal of the propensity score analysis is to mitigate bias by adjusting for confounders across groups, thereby isolating the true effect of the treatment on the outcome. This approach assumes that the treatment is not randomly assigned; i.e., whether a stock is chosen for short selling by market makers is influenced by the stock's observable covariates.[26] Recent studies have explored sophisticated machine-learning techniques, beyond traditional logistic regression analysis, to estimate propensity scores by capturing nonlinear covariate relationships.[27]

Propensity score weighting (PSW) is a statistical technique designed to address confounding bias by applying weights based on the inverse of the estimated propensity scores. It adjusts for selection bias by ensuring that pre-treatment covariates are evenly distributed between the treatment and control groups (Morgan and Todd, 2008). Thus, PSW aims to balance the distribution of confounders across groups, thereby approximating randomization and reducing bias. Propensity score matching (PSM) is another popular method to control for confounding bias. It matches samples in the treatment group with those in the control group based on similar propensity scores to estimate the ATE on the treated (Imbens, 2004). A major drawback of PSM is the potential loss of unmatched observations, which can decrease the sample size and potentially limit the generalizability of the findings. In contrast, PSW retains all samples in the analysis, thereby preserving the original number of observations and potentially enhancing the study's external validity. By assigning

[26] Nevertheless, this approach can only adjust for confounders that are measured and included in the analysis, leaving the possibility of residual confounding bias from unmeasured factors.

[27] Examples include Random Forest (Lee, Lessler, and Stuart 2010), recursive partitioning or tree-based methods (Lee, Lessler, and Stuart, 2010; Setoguchi et al., 2008), neural networks (Setoguchi et al., 2008), and bagging or boosting (Lee, Lessler, and Stuart, 2010; McCaffrey, Ridgeway, and Morral, 2004; McCaffrey et al., 2013).

weights based on propensity scores, this method adjusts for confounding without discarding data, making it more efficient than matching.

Recent methods have been developed to address the limitations of conventional propensity-score weighting, which can result in excessively high or low weights. In particular, overlap propensity-score weighting addresses issues arising from poor overlap between the treatment and control groups by limiting weights to the minimum propensity score within the range of common support (Li, Morgan, and Zaslavsky, 2017). As shown by Thomas, Li, and Pencina (2020), overlap propensity-score weighting reduces the influence of extreme scores and enhances the stability and accuracy of causal estimates, improving the robustness of causal inference. We implement the overlap weighting analysis using the PSweight package (Li, Morgan, and Zaslavsky, 2017).

4. Empirical Results

This section presents the results of our empirical analysis based on the methodologies described in Section 3, focusing on the impact of SSFs market-makers' short selling on the quality of the futures market and its cross-market quality with the underlying spot markets. Through these results, our primary objective is to determine whether SSFs market-makers' short selling represents a beneficial mechanism for overall market functioning or harmful behavior.

4.1. Clustering Short-Selling Activities: x

We perform clustering on the short-selling ratio using the DBSCAN algorithm.[28] The short-selling ratio is defined as the ratio of short selling volume by SSFs market makers to the total trading volume of each underlying stock on a given trading day. The clustering process results in three distinct groups, formed around thresholds of approximately 0.44% and 0.53%, which closely align with the internal guideline of 0.5%. One cluster, which exceeded the 0.53% threshold, is identified as an outlier and labeled *Aggressive*, as this cluster represents samples which engaged in aggressive short selling beyond the 0.5% guideline. The remaining two clusters, both below the 0.53% threshold, are categorized as *Compliant* groups. Among these, the cluster between the 0.44% and 0.53% thresholds is labeled *Reluctantly Compliant*, reflecting

[28]The DBSCAN algorithm controls clustering behavior through a pair of primary hyperparameters: *min samples*, which specifies the minimum number of points required to form a dense region, and *epsilon*, which defines the radius for neighborhood points. In our implementation, these were configured as (*min samples, epsilon*) = $(50, 10^{-4})$.

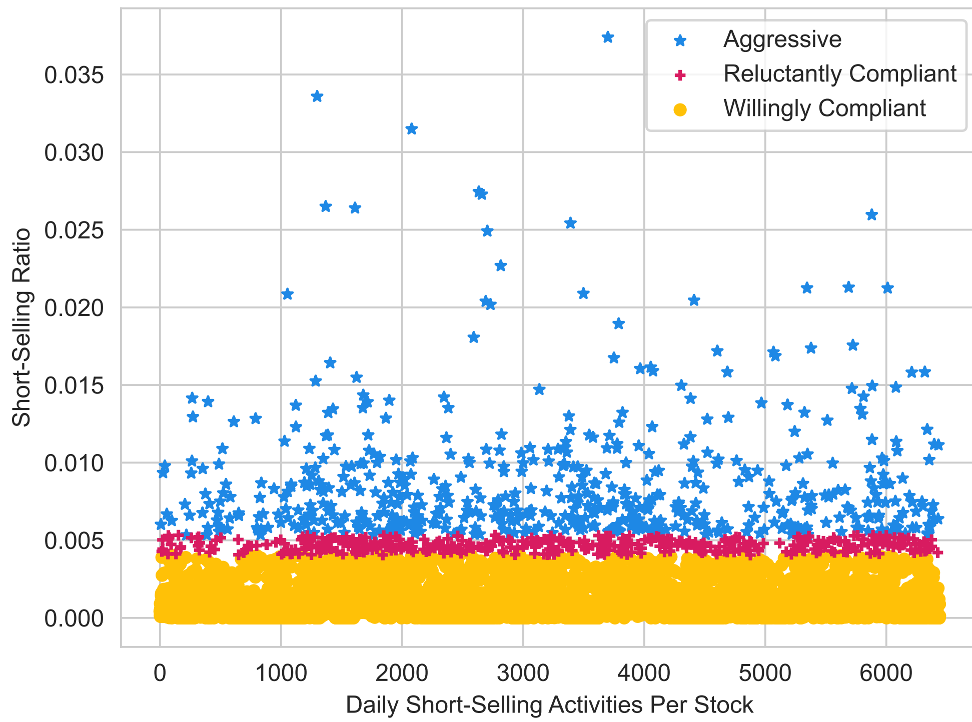

Figure 1. Clustering of Short-Selling Ratios with DBSCAN. This figure presents the DBSCAN clustering results of daily short-selling activity by SSFs market makers, classified by underlying stock based on the short-selling ratio. The clusters are visually distinguished using different markers: '★' represents aggressive clusters, '+' indicates reluctantly compliant clusters, and '●' denotes willingly compliant clusters. The threshold between aggressive and reluctantly compliant clusters is set at 0.53%, and between reluctantly and willingly compliant clusters at 0.44%.

those who would have preferred to take a bigger short position but adhered to the 0.5% guideline. In contrast, the cluster with activity below the 0.44% threshold is labeled *Willingly Compliant*, representing voluntarily limited short-selling activities, independent of the 0.5% short-selling regulation. Figure 1 visualizes the clustering results, where the clusters are clearly delineated into three groups based on these thresholds, with aggressive short selling marked with '★,' reluctantly compliant short selling marked with '+,' and willingly compliant short selling marked with '●.'

4.2. Sample Analysis: y

As illustrated in Section 3, the market quality variables are grouped into three categories: liquidity, volatility, and price efficiency. Table 1 presents the mean differences and their statistical significances across clusters for the variables used in the analysis, based on a sample of 6,441 observations; see Table A.1 in the Appendix for the complete descriptive statistics of our dataset. Panel A reports the differences in the sample means of market quality variables across clusters, while Panel B presents the key factor variable differences used in the supervised machine-learning classification analysis to calculate propensity scores.

Panel A of Table 1 reveals that the *Aggressive* cluster generally exhibited better market quality compared to the *Compliant* clusters in the categories of *volatility* and *price efficiency*, while the mean differences were less pronounced in the *liquidity* category. For instance, the sample means in *realized volatility* were lower in the *Aggressive* cluster than the *Compliant* clusters, regardless of the minutes that we used for calculation. This result also occurred for the sample means for the *variance ratio* and *execution shortfall*, further suggesting the superior contribution of the *Aggressive* cluster to market quality compared to the *Compliant* clusters. When comparing the *Reluctantly Compliant* cluster to the *Willingly Compliant* cluster, improvements in most market quality variables were less pronounced or exhibited unclear statistical significance. It is important to note, however, that these sample statistics do not address potential endogeneity issues.

Panel B of Table 1 highlights the key factors potentially influencing SSFs market-makers' short-selling intensity. The difference in the sample means of *market basis gap* between the *Aggressive* and *Compliant* clusters was negative with statistical significance, indicating that deeper backwardation was associated with aggressive short selling. Figure 2 further supports this conjecture, showing the *Aggressive* cluster skewed toward negative values, while the *Compliant* clusters centered closer to zero and positive values. Additionally, the *sector* dummy variable showed a higher mean value in the *Aggressive* cluster than in the *Compliant* clusters, implying a greater likelihood of short selling in the financial sector than in non-financial sectors. These observations indicate a meaningful distinction among the clusters and support their usefulness in estimating the propensity of short-selling intensity.

Table 1. Difference of Sample Means across Clusters. This table presents the mean differences and their statistical significances across clusters for the variables used in the analysis, based on a sample of 6,441 observations. Three clusters are categorized: *Aggressive* (AG), *Reluctantly Compliant* (RC), and *Willingly Compliant* (WC). The samples are winsorized at the 1^{st} and 99^{th} percentiles to mitigate the effect of outliers. The t-statistics compare the mean differences between AG and RC and between RC and WC. Panel A displays the mean differences in market quality variables across clusters. *Effective spread*, *price impact*, *execution shortfall*, and *pricing error variability* are scaled by a factor of 10,000. Similarly, *Amihud illiquidity ratio* is scaled by a factor of one billion. *Execution shortfall* and *pricing error variability* are log-transformed for analysis. Panel B shows the mean differences in variables used for the propensity score calculation in overlap weighting. Statistical significance levels are denoted by: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Category	Variable	AG – (RC ∪ WC)	RC – WC
Panel A. Market Quality Variables			
Liquidity	Effective Spread	-1.5015* (-1.8645)	-2.5287** (-2.5672)
	Price Impact	-0.0044 (-0.0387)	-0.1174 (-1.0818)
	Amihud Illiquidity Ratio	1.3274 (0.6350)	-2.3554 (-1.3303)
Volatility	1-Min Realized Volatility	-0.0596** (-2.4505)	0.0403 (1.5109)
	5-Min Realized Volatility	-0.0642*** (-2.8840)	0.0399 (1.6378)
	10-Min Realized Volatility	-0.0585*** (-2.7172)	0.0454* (1.9077)
Price Efficiency	Variance Ratio	-5.0862** (-2.1373)	-2.5760 (-0.7046)
	Execution Shortfall	-0.1293*** (-3.7645)	0.0927** (2.2817)
	Pricing Error Variability	-0.4393 (-1.6962)	0.3473** (1.918)
Panel B. Variables for Propensity Score Calculation			
	Ask Market Depth	-0.0135 (-1.3049)	-0.0691*** (-5.9489)
	Bid Market Depth	0.0108 (1.2817)	-0.0372*** (-3.7705)
	Market Makers' Ask Trading Amount	-0.0084 (-1.1053)	-0.0277*** (-3.4314)
	Market Makers' Ask Trading Profit	304,235*** (6.0483)	584,463*** (8.2954)
	Market Makers' Bid Trading Amount	0.0378*** (5.5248)	0.0379*** (4.9701)
	Market Makers' Bid Trading Profit	-231,082*** (-9.2583)	-376,903*** (-9.0559)
	Market Basis Gap	-0.5209*** (-9.4595)	-0.6376*** (-10.5153)
	Max Price Change	-0.2742*** (-3.4635)	-0.1784 (-1.6343)
	Trading Amount	-4,856*** (-3.2836)	1,134 (0.4938)
	Trading Balance	-220 (-1.4591)	-916*** (-3.6217)
	Sector	0.1108*** (5.7611)	0.0481** (2.3219)
	Short-Selling Balance	0.0016*** (5.6319)	0.0012*** (3.7952)
	Stock Trading Amount	-198,838*** (-4.6413)	107,815 (1.0570)
	Stock Volatility	-0.2936*** (-3.3793)	-0.0543 (-0.4447)

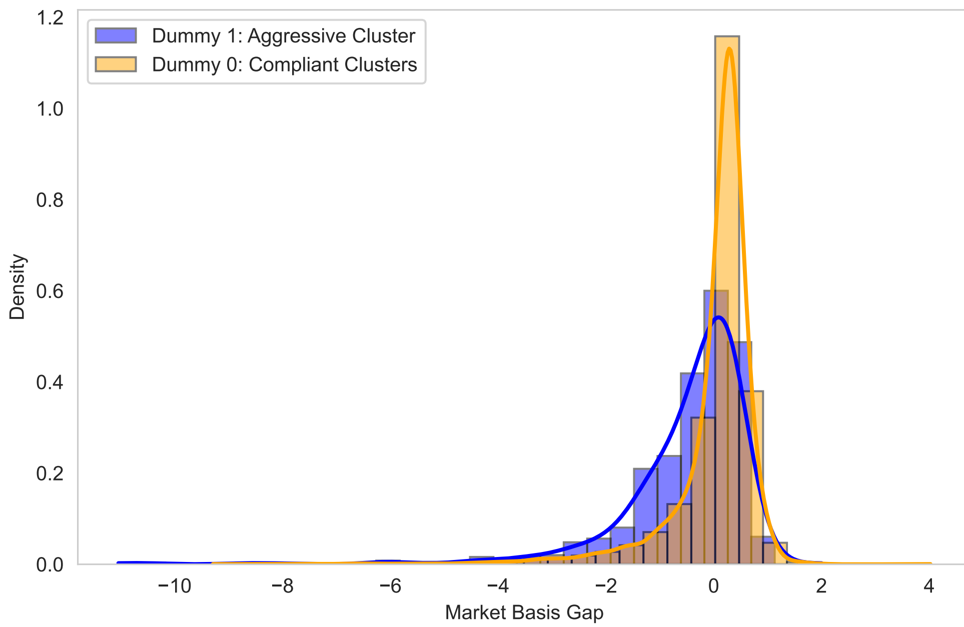

Figure 2. Distribution of *Market Basis Gap*. This figure compares the distribution of *market basis gap* for *Aggressive* (dummy = 1) and *Compliant* (dummy = 0) short-selling activities. The *Aggressive* cluster exhibits higher density in the backwardation region (i.e., negative gap), suggesting that deeper backwardation is associated with more aggressive short-selling behavior. This highlights *market basis gap* as a key variable in identifying short-selling propensities.

4.3. Propensity Score Estimation: $f(\cdot)$

Four supervised machine-learning classification algorithms were employed to calculate propensity scores: GBM,[29] RF,[30] XGB,[31] and LR. In selecting the variables for propensity score calculation, we first performed univariate logistic regressions and retained the variables with statistically significant contributions, defined as those with p-values less than 0.1. Then, we assessed the pairwise correlation among the selected variables and excluded those exhibiting high correlations, specifically those with an absolute correlation coefficient greater than 0.5, to prevent potential multicollinearity.

[29]GBM is an ensemble machine-learning classification method that builds models sequentially, with each model correcting the errors of its predecessor by minimizing a specified loss function using gradient descent (Friedman, 2001). We implemented GBM using the *scikit-learn* library in Python (Pedregosa et al., 2011.)

[30]RF is an ensemble machine-learning classification method that constructs multiple decision trees and combines their predictions to improve accuracy and reduce overfitting by introducing randomness in both feature selection and data sampling (Breiman, 2001). We implemented RF using the *scikit-learn* library in Python (Pedregosa et al., 2011).

[31]XGB is an efficient and scalable implementation of the gradient boosting framework, incorporating techniques such as parallel processing, regularization, and tree pruning to enhance performance and prevent overfitting. We implemented XGB using the *xgboost* library in Python (Chen and Guestrin, 2016).

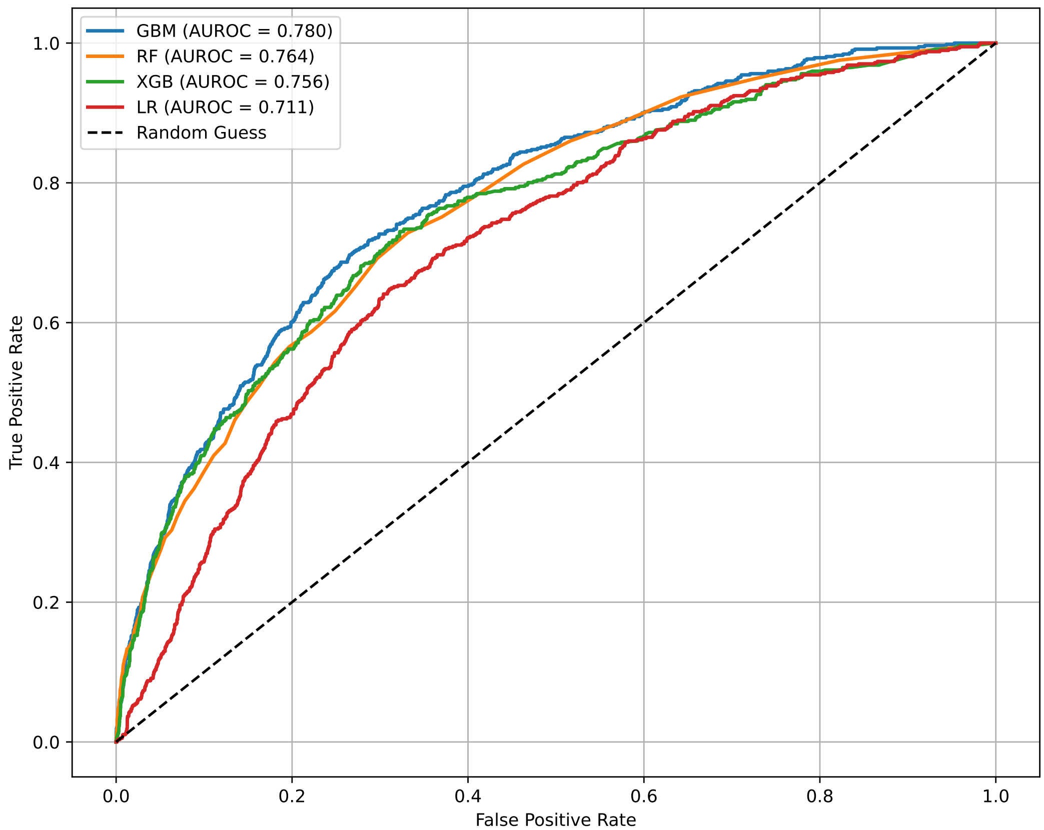

Figure 3. AUROC Curves for Classification Models. This figure presents the Area Under the Receiver Operating Characteristic (AUROC) curves for four classification models: Gradient Boosting Machine (GBM), Random Forest (RF), eXtreme Gradient Boosting (XGB), and Logistic Regression (LR). The AUROC curve plots the True Positive Rate against the False Positive Rate, measuring each model's ability to distinguish between classes across various classification thresholds. GBM achieved the highest AUROC (0.780), followed by RF (0.764), XGB (0.756), and LR (0.711). The dashed diagonal line represents random guessing (AUROC = 0.5).

Table 2. *K*-Fold Propensity Score Estimation (*K*=5). This table compares models used to estimate propensity scores for two categories of short selling: *Aggressive* and the combined group of *Reluctantly Compliant* and *Willingly Compliant* clusters. The AUROC values indicate out-of-sample performance based on Receiver Operating Characteristic (ROC) analysis. The 99% confidence intervals are constructed using the method proposed by DeLong, DeLong, and Clarke-Pearson (1988). The rightmost column reports AUROC differences from the Logistic Regression (LR) baseline, along with associated Z-statistics in parentheses. All models are evaluated using five-fold cross-validation with AUROC computed from out-of-fold (OOF) predictions. Statistical significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Model	AUROC	99% C.I.	Difference from LR (Z-statistics)
GBM	0.780	[0.754, 0.806]	0.069^{***} (8.6814)
RF	0.764	[0.738, 0.790]	0.053^{***} (5.7260)
XGB	0.756	[0.728, 0.784]	0.045^{***} (5.0507)
LR	0.711	[0.683, 0.738]	N/A

To enhance robustness and accuracy in model selection while reducing the risk of overfitting, we performed K-fold cross-validation with $K = 5$. The out-of-fold (OOF) predictions were aggregated across all test sets to generate a comprehensive OOF prediction dataset. The Area Under the Receiver Operating Characteristic (AUROC) curve serves as a critical metric for evaluating the performance of classification models, reflecting the model's ability to differentiate between classes across various threshold levels. An AUROC value close to 1 indicates superior model performance, while a value of 0.5 suggests performance equivalent to random guessing (Hanley and McNeil, 1982).[32] Figure 3 visualizes the AUROC curves for the models and Table 2 presents the out-of-sample AUROC results of the estimated models, demonstrating that the GBM model outperforms the others.

Given that LR is a standard method for propensity score estimation, we conducted a DeLong's test (DeLong, DeLong, and Clarke-Pearson, 1988) to evaluate whether the difference in AUROC between LR and the alternative models was statistically significant. As reported in Table 2, our results confirmed that the GBM model significantly outperformed LR in propensity score estimation. This superiority stems from the fact that advanced machine-learning models can capture both linear and nonlinear relationships, unlike LR, which is limited to linear associations. This reveals that machine-learning models account for complex variable interactions, improving out-of-sample performance for propensity score estimation.

[32]The AUROC intuitively represents the probability that a randomly selected 'positive' sample ranks higher than a 'negative' sample based on the estimated propensity scores out-of-sample. Its statistical interpretation aligns with the Wilcoxon rank test (Hanley and McNeil, 1982) and is closely related to the Gini coefficient (Breiman et al., 2017).

Table A.2 in the Appendix displays the "feature importance"[33] used in calculating the propensity scores. The *market basis gap* variable was a critical determinant in the model, capturing the excessive deviation between the futures price and the spot price from its theoretically fair level. Negative market basis gap values indicate deeper backwardation, which is strongly associated with an increased likelihood of short-selling activity. When the backwardation becomes more pronounced, SSFs market makers often engage in increased short selling of stocks to hedge their long futures positions, taking advantage of the mispricing between the futures and spot markets to enhance profitability. It is also noteworthy that stocks in the financial sector are more easily targeted for short selling, *ceteris paribus*, due to their high leverage, sensitivity to economic conditions, and transparency from regulatory reporting, which makes vulnerabilities more apparent. They face systemic risks, reputational sensitivities, and interest rate fluctuations, making them more volatile and attractive to market makers. These factors, combined with the sector's reliance on trust and susceptibility to negative sentiment, make stocks in the financial sector prime targets during periods of uncertainty or stress. These characteristics reflect the importance of observable covariates in influencing market-makers' short-selling decisions, aligning with Section 3.5.2. Figure A.2 in the Appendix supports our reasoning, showing that a negative *market basis gap* (backwardation) is associated with aggressive short selling, while the *sector* indicator variables, which take the value 1 for the financial sector, also increase the likelihood of SSFs market makers' aggressive short-selling activities.

4.4. Overlap Propensity-Score Weighting: $y = f(x)$

Our empirical analysis exploits the binary structure of the overlap propensity-score weighting framework, which facilitates two distinct comparison levels: (i) *Aggressive* versus *Compliant* (including both *Reluctantly* and *Willingly Compliant*) short-selling activities and (ii) *Reluctantly Compliant* versus *Willingly Compliant* short-selling behaviors. This framework allows us to evaluate the causal effects of market-makers' short-selling activities on various market quality dimensions, offering valuable insights into the regulation of short-selling practices.

Although GBM was selected as the optimal model based on its out-of-sample performance, we subsequently retrained it using the full in-sample dataset to perform the final empirical analysis and to provide a comprehensive representation of patterns

[33]Feature importance in machine learning refers to evaluating input variables based on their impact on a model's predictions, highlighting the data's most influential attributes to improve accuracy and understand the key drivers of the results.

within the study period. This approach is consistent with standard practices in the existing literature and with our robustness analysis based on panel regressions in Section 4.5.

4.4.1. The effects of market-makers' short selling on market quality

Table 3 summarizes the results of the overlap weighting analyses, providing the ATE estimates for the market quality metrics. The first comparison focused on *Aggressive* versus *Compliant* short selling. The results demonstrate that aggressive short selling by SSFs market makers significantly enhanced market quality across all dimensions. Specifically, liquidity improved, as evidenced by a significant reduction in *effective spread*, suggesting better trading conditions. *Amihud illiquidity ratio* was not statistically significant; however, this was expected given its conceptual role as a proxy for price impact rather than a direct measure of liquidity (Brauneis et al., 2021). In our analysis, both *Amihud illiquidity ratio* and *price impact* showed similar signs and statistical insignificance.[34] These findings ultimately reinforce that *effective spread* is the most critical liquidity measures in this context. *Realized volatility* exhibited a statistically significant reduction across all time intervals, reflecting its stabilizing effect on market fluctuations. Price efficiency also improved markedly, as evidenced by significant enhancements in *variance ratio* and *execution shortfall*, collectively highlighting a substantial increase in market efficiency.

The second comparison examined the binary comparison between the *Reluctantly Compliant* and *Willingly Compliant* clusters. The results indicate that short selling in the *Reluctantly Compliant* cluster did not significantly improve any aspect of market quality. Measures of liquidity showed no meaningful improvement, and other dimensions such as volatility and price efficiency also remained largely unaffected. These findings suggest that the restrictive nature of short selling in the *Reluctantly Compliant* cluster limited its overall market impact.

Overall, the comparison between the *Aggressive* and *Compliant* clusters indicates that the *Aggressive* cluster's short selling contributed to improvements in most measures of liquidity, volatility, and price efficiency. However, within the *Compliant* category, the *Reluctantly Compliant* cluster's short selling did not significantly improve overall market quality compared to the *Willingly Compliant* cluster. This

[34]Our results based on *Amihud illiquidity ratio* may differ from those observed in stock markets. For example, while *Amihud illiquidity ratio* conceptually considers the relationship between volume and liquidity to be negative, in the KRX futures market, this relationship shows almost no correlation (-0.028), highlighting the possibility of distinct results between stock and futures markets due to structural differences.

Table 3. Overlap Propensity-Score Weighting Analysis. This table presents the impact of short selling (dummy1) on liquidity, volatility, and price efficiency-related variables. Three clusters are categorized: *Aggressive* (AG), *Reluctantly Compliant* (RC), and *Willingly Compliant* (WC). For the 'AG vs. RC ∪ WC' comparison, dummy1 represents aggressive short selling, while for 'RC vs. WC', dummy1 represents reluctantly compliant short selling. The overlap weighting method reports the Average Treatment Effect (ATE) with the associated *p*-value in parentheses. More negative values indicate better market quality. Statistical significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variable	AG vs. RC ∪ WC	RC vs. WC
Panel A. Liquidity		
Effective Spread	-1.7825* (0.0507)	-1.8827 (0.1009)
Price Impact	0.0704 (0.5382)	-0.1660 (0.1841)
Amihud Illiquidity Ratio	1.4865 (0.5400)	-3.1670 (0.1121)
Panel B. Volatility		
1-Min Volatility	-0.0450* (0.0780)	-0.0158 (0.5855)
5-Min Volatility	-0.0451* (0.0584)	-0.0174 (0.5177)
10-Min Volatility	-0.0408* (0.0771)	-0.0092 (0.7300)
Panel C. Price Efficiency		
Variance Ratio	-7.3103** (0.0272)	-1.4471 (0.7322)
Execution Shortfall	-0.1181*** (0.0009)	-0.0311 (0.4703)
Pricing Error Variability	-0.2620 (0.2741)	0.1239 (0.5122)

suggests that the internal guideline of 0.5% imposed constraints on short-selling activities, thereby limiting the ability of market makers to fully perform their role.[35]

4.4.2. Cross-market analysis

One of the key functions of SSFs market makers is to reduce basis discrepancies between the futures and spot markets, thereby facilitating price convergence. To assess whether SSFs market-makers' short selling played a significant role in this function, we implemented overlap propensity-score weighting between the *Aggressive* and *Compliant* clusters using the *market basis gap* variable, defined in Section 3.4.4, as a cross-market quality measure.

As shown in Figure 4, our findings indicate that market-makers' short-selling activities did not significantly reduce *market basis gap*, suggesting that they did not meaningfully contribute to cross-market quality in price discovery. The box plots show nearly identical distributions between the two groups, with the ATE differential

[35]Market makers were required to comply with the 0.5% guideline to avoid potential disadvantages in future market-making activities. According to market-making regulations, market makers who do not agree with the exchange's future market-making operational plans during the re-selection process may be excluded.

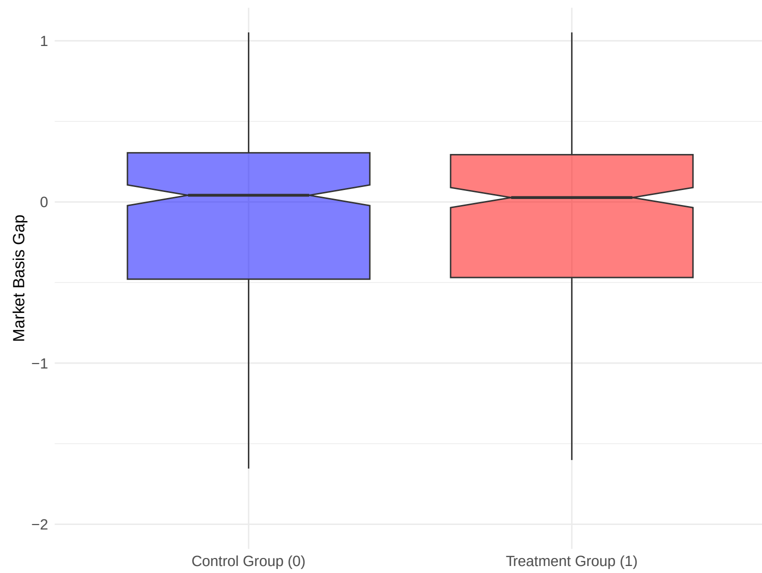

Figure 4. Adjusted Distributions of *Market Basis Gap* Using Overlap Propensity-Score Weighting. This box plot illustrates the distribution of *market basis gap* for the *Compliant* short-selling cluster (Control Group, 0) and the *Aggressive* short-selling cluster (Treatment Group, 1), based on overlap propensity-score weighting. The overlap weights adjust each data point's contribution to the statistical calculations—values with higher weights have greater influence on medians and quartiles, while those with lower weights have reduced impact. The estimated Average Treatment Effect (ATE) between the two groups is -0.0009 with a p-value of 0.9832, indicating no statistically significant difference.

of -0.0009 and its p-value of 0.9832, further supporting this result. Our interpretation of this result is that the stringent short-selling restrictions prevented market makers from effectively fulfilling this role, thereby limiting their ability to arbitrage inefficiencies between the two markets.

4.5. Robustness Analyses: Panel Regressions

While propensity score weighting balances observable covariates, it relies on the correct specification of the propensity score model. Panel regression complements this by directly modeling the relationship between covariates and outcomes while accounting for unobserved heterogeneity across samples. Accordingly, our robustness analysis incorporates the panel regression in two ways: '*with dummies*' and '*without dummies*,' capturing the nuanced relationships and variations that the estimated propensity scores cannot explicitly address. This approach offers deeper insights into the effects of continuous predictors, providing a more comprehensive understanding of the underlying interactions.

In the '*with dummies*' approach, we categorize short-selling activity clusters as outlined in Section 4.4, analyzing two cases: (i) *Aggressive* versus *Compliant* (including both *Reluctantly* and *Willingly Compliant*), where *Aggressive* short selling is assigned $D = 1$; and (ii) *Reluctantly Compliant* versus *Willingly Compliant*, where *Reluctantly Compliant* short selling is assigned $D = 1$. In contrast, the '*without dummy*' approach examines how market quality coefficients respond to variations in short-selling intensity. To capture this, we use short-selling ratio variables, with the short-selling ratio defined as short-selling volume divided by total trading volume, allowing for a more granular analysis beyond binary classifications, although it may be subject to potential confounding bias.

4.5.1. Panel regression with dummies

We compared Table 3's results with those obtained from panel regression analysis. The key benefit of using panel-data analysis is that it offers a detailed view of market quality by incorporating aspects of market-microstructure models. Following the methodologies outlined by Eom, Ok, and Park (2007) and Bryzgalova, Pavlova, and Sikorskaya (2023), we conducted the panel-data analysis separately for each market quality variable is analyzed using the following panel regression specification for each SSFs contract i on day t:

$$y_{it} = \beta_0 + \gamma_D D_{it} + \delta' C_{it} + \alpha_i + \mu_t + \varepsilon_{it},$$

where D_{it} is a dummy variable defined under two different settings: (i) $D_{it} = 1$ indicates *Aggressive* short selling where $D_{it} = 0$ corresponds with *Compliant* short selling (including both *Reluctantly* and *Willingly Compliant*); and (ii) $D_{it} = 1$ represents *Reluctantly Compliant* short selling where $D_{it} = 0$ denotes *Willingly Compliant* short selling, distinguishing between the two *Compliant* subgroups. Furthermore, β_0 indicates the constant term, capturing the dependent variable baseline level, C_{it} represents the control variables, including the variables used in the propensity score estimation, along with the log-transformed price, and $(\alpha_i, \mu_t, \varepsilon_{it})$ denote firm-specific effects, time-specific effects, and an independent error term with a mean of zero, respectively.

Table 4 presents the panel regression results, summarizing the coefficients for the dummy1 variable. While the statistical significance of certain variables was reduced compared to the overlap weighting analysis, the results remained consistent regarding the effect direction, thereby corroborating the findings of the overlap weighting approach. When comparing *Aggressive* and *Compliant* short selling, the panel regression results reaffirmed that market-makers' *Aggressive* short selling positively

Table 4. Panel Regression Analysis with Dummies. This table presents the impact of short selling (dummy1) on liquidity, volatility, and price efficiency-related variables. Three clusters are categorized: *Aggressive* (AG), *Reluctantly Compliant* (RC), and *Willingly Compliant* (WC). For the 'AG vs. RC ∪ WC' comparison, dummy1 represents aggressive short selling, while for 'RC vs. WC', dummy1 represents reluctantly compliant short selling. The panel regression estimates the coefficient ($\widehat{\gamma_D}$) for dummy1, with the corresponding p-value reported in parentheses. The regression includes a constant term, which is omitted for brevity. More negative values indicate better market quality. Statistical significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variable	AG vs. RC ∪ WC	RC vs. WC
Panel A. Liquidity		
Effective Spread	-1.1921 (0.2883)	-1.1275 (0.3711)
Price Impact	0.0152 (0.8982)	-0.1325 (0.2466)
Amihud Illiquidity Ratio	1.8090 (0.2985)	-3.1976* (0.0510)
Panel B. Volatility		
1-Min Realized Volatility	-0.0249 (0.1957)	-0.0015 (0.9528)
5-Min Realized Volatility	-0.0360* (0.0529)	-0.0090 (0.6878)
10-Min Realized Volatility	-0.0401** (0.0187)	-0.0044 (0.8415)
Panel C. Price Efficiency		
Variance Ratio	-8.1848** (0.0317)	-1.7613 (0.6622)
Execution Shortfall	-0.0299 (0.3892)	0.0252 (0.2466)
Pricing Error Variability	-0.4376* (0.0958)	-0.0877 (0.6276)

Table 5. Panel Regression Results by Market Quality Variables without Dummies. This table presents the impact of market makers' short-selling ratio on liquidity, volatility, and price efficiency-related variables. The panel regression reports the estimated coefficient ($\widehat{\gamma}_M$) and its corresponding p-value in parentheses. A constant term is included in the regression but omitted from the table for brevity. More negative values indicate better market quality. Statistical significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Variable	**Estimated Coefficient** ($\widehat{\gamma}_M$)
Panel A. Liquidity	
Effective Spread	-201.7368 (0.1200)
Price Impact	-4.0569 (0.7496)
Amihud Illiquidity Ratio	-8.6928 (0.9703)
Panel B. Volatility	
1-Min Realized Volatility	-2.6710 (0.2060)
5-Min Realized Volatility	-3.9040* (0.0525)
10-Min Realized Volatility	-3.8428** (0.0465)
Panel C. Price Efficiency	
Variance Ratio	-823.7770** (0.0458)
Execution Shortfall	-7.4872* (0.0631)
Pricing Error Variability	-65.4314* (0.0506)

contributed to market quality. In contrast, comparing *Reluctantly Compliant* and *Willingly Compliant* short selling revealed more limited effects.

4.5.2. Panel regression without dummies

To verify the previous results, we examine the impact of market-makers' short-selling ratio on market quality using a panel regression specification, excluding the short-selling cluster dummy variables. The analysis follows the same framework as described in Section 4.5.1, replacing the dummy variable with the short-selling ratio as the independent variable, where the modified specification for each SSFs contract i on day t is expressed as

$$y_{it} = \beta_0 + \gamma_M M_{it} + \delta' C_{it} + \alpha_i + \mu_t + \varepsilon_{it},$$

where M_{it} represents the market maker's short-selling ratio.

Table 5 provides the results of the panel regression analysis, focusing on the impact of market-makers' short-selling ratio on market quality. The table includes the estimated coefficients for the short-selling ratio variable, highlighting their statistical significance and effects across various market quality dimensions. Consistent with the results presented in Table 4, we found that higher levels of short selling are associated with improved market quality.

In terms of liquidity, while none of the variables showed statistical significance, the negative coefficients indicate an improvement in liquidity. For volatility, *realized volatility* decreased significantly at the five- and ten-minute intervals, indicating a stabilizing effect on market fluctuations. Regarding price efficiency, all three metrics—*variance ratio*, *execution shortfall*, and *pricing error variability*—demonstrated statistically significant improvements, reinforcing the conclusion that market-makers' short selling enhanced price efficiency. These findings support the robustness of our primary analysis and further validate the positive role of market-makers' short selling in improving overall market quality.

5. Conclusion

This paper provides a novel contribution to a better understanding of the roles of short selling and market makers. By leveraging advanced machine-learning techniques and the overlap weighting of propensity scores to control for confounding selection bias, our analysis demonstrates that the short-selling activities of SSFs market makers significantly improved the SSFs market quality by increasing liquidity, reducing volatility and enhancing price efficiency. The application of both overlap weighting and conventional panel regression analysis confirms the positive influence of short selling on market quality and price efficiency within the SSFs market. It is remarkable that the SSFs market-makers' short-selling activities had a favorable impact, despite the stringent restrictions imposed on short selling such as the implicit KRX 0.5% guideline. However, it is also noteworthy that these benefits did not extend to the spot markets, as evidenced by the persistence of the backwardation effect, which indicates a limited impact of short selling on the underlying stock markets. This discrepancy indicates that the guideline's short-selling restrictions on market makers may not fully address cross-market inefficiencies or anomalies.

Our findings offer valuable insights into market-makers' short-selling role in promoting liquidity and market efficiency, particularly under restrictive trading-rule conditions. This paper is the first to examine the role of SSFs market makers under restrictive short-selling conditions, providing new insights into the short-selling debate. The findings are expected to inform future regulatory policies by providing cross-market perspectives on the effectiveness of short selling and its impact on market quality.

Future research would greatly benefit from a more granular analysis of market makers' strategies and profitability, offering more profound insights into their short-selling motivations. In addition, extending the analysis to include single-stock options (SSOs) markets could provide a broader perspective on the effects of market-making strategies across different financial instruments. Such extensions would provide further insights into how market-makers' short-selling practices influence various market segments. In particular, examining SSOs in this context could contribute to a more comprehensive understanding of intermarket dynamics.

References

Aggarwal N, Thomas S, 2019. When stock futures dominate price discovery. *J. Futures Markets* 39(3):263–278.

Amihud Y, 2002. Illiquidity and stock returns: Cross-section and time-series effects. *J. Financial Markets* 5(1):31–56.

Andersen TG, Bollerslev T, Diebold FX, Labys P, 2003. Modeling and forecasting realized volatility. *Econometrica* 71(2):579–625.

Aoyagi J, Ito Y, 2024. Coexisting exchange platforms: Limit order books and automated market makers. *J. Political Econom. Microeconom.* Forthcoming.

Asquith P, Pathak PA, Ritter JR, 2005. Short interest, institutional ownership, and stock returns. *J. Financial Econom.* 78(2):243–276.

Atmaz A, Basak S, Ruan F, 2024. Dynamic equilibrium with costly short-selling and lending market. *Rev. Financial Stud.* 37(2):444–506.

Battalio R, Schultz P, 2011. Regulatory uncertainty and market liquidity: The 2008 short sale ban's impact on equity option markets. *J. Finance* 66(6):2013–2053.

Beber A, Pagano M, 2013. Short-selling bans around the world: Evidence from the 2007–09 crisis. *J. Finance* 68(1):343–381.

Beneish MD, Lee CMC, Nichols DC, 2015. In short supply: Short-sellers and stock returns. *J. Accounting Econom.* 60(2):33–57.

Bessembinder H, Hao J, Zheng K, 2015. Market making contracts, firm value, and the IPO decision. *J. Finance* 70(5):1997–2028.

Bessembinder H, Hao J, Zheng K, 2020. Liquidity provision contracts and market quality. *Rev. Financial Stud.* 33(1):44–74.

Black F, 1975. Fact and fantasy in the use of options. *Financial Analysts J.* 31(4):36–41, 61–72.

Blau BM, Whitby RJ, 2018. How does short selling affect liquidity in financial markets? *Finance Res. Lett.* 25(C):244–250.

Boehmer E, Jones CM, Wu Z, Zhang X, 2020. What do short sellers know? *Rev. Finance* 24(6):1203–1235.

Boehmer E, Jones CM, Zhang X, 2008. Which shorts are informed? *J. Finance* 63(2):491–527.

Boehmer E, Jones CM, Zhang X, 2013. Shaking short sellers: The 2008 shorting ban. *Rev. Financial Stud.* 26(6):1363–1400.

Boehmer E, Wu Z, 2013. Short selling and the price discovery process. *Rev. Financial Stud.* 26(2):287–322.

Brauneis A, Mestel R, Riordan R, Theissen E, 2021. How to measure the liquidity of cryptocurrency markets? *J. Bank. Finance* 124:106041.

Breiman L, 2001. Random forests. *Mach. Learn.* 45(1):5–32.

Breiman L, Friedman J, Olshen RA, Stone CJ, 2017. *Classification and Regression Trees* (Routledge).

Brogaard J, Hendershott T, Riordan R, 2017. High frequency trading and the 2008 short-sale ban. *J. Financial Econom.* 124(1):22–42.

Brunnermeier MK, Oehmke M, 2014. Predatory short selling. *Rev. Finance* 18(6):2153–2195.

Brunnermeier MK, Pedersen LH, 2005. Predatory trading. *J. Finance* 60(4):1825–1863.

Bryzgalova S, Pavlova A, Sikorskaya T, 2023. Retail trading in options and the rise of the big three wholesalers. *J. Finance* 78(6):3465–3514.

Chang EC, Cheng JW, Yu Y, 2007. Short-sales constraints and price discovery: Evidence from the Hong Kong market. *J. Finance* 62(5):2097–2121.

Chen T, Guestrin C, 2016. Xgboost: A scalable tree boosting system. *Proc. 22nd ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining*, 785–794.

Choe H, Lee H-J, 2012. Short sales prohibition: Different effects on market efficiency. *Asian Rev. Financial Res.* 25(2):161–202, (in Korean).

Clark-Joseph AD, Ye M, Zi C, 2017. Designated market makers still matter: Evidence from two natural experiments. *J. Financial Econom.* 126(3):652–667.

Comerton-Forde C, Jones CM, Putnins TJ, 2016. Shorting at close range: A tale of two types. *J. Financial Econom.* 121(3):546–568.

Coughlan J, Orlov AG, 2023. High-frequency trading and market quality: Evidence from account-level futures data. *J. Futures Markets* 43(8):1126–1160.

Danielsen B, Sorescu S, 2001. Why do option introductions depress stock prices? An empirical study of diminishing short sale constraints. *J. Financial Quant. Anal.* 36(4):451–484.

Daures-Lescourret L, Monias S, 2023. Fragmentation and strategic market-making. *J. Financial Quant. Anal.* 58(4):1675–1700.

Dechow PM, Hutton AP, Meulbroek L, Sloan RG, 2001. Short-sellers, fundamental analysis, and stock returns. *J. Financial Econom.* 61(1):77–106.

DeLong ER, DeLong DM, Clarke-Pearson DL, 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 44(3):837–845.

Demsetz H, 1968. The cost of transacting. *Q. J. Econom.* 82(1):33–53.

Deutsche Börse AG, 2023. Designated Sponsor Guide, version 13.1. Technical report, Deutsche Börse AG.

Diamond DW, Verrecchia RE, 1987. Constraints on short-selling and asset price adjustment to private information. *J. Financial Econom.* 18(2):277–311.

Diether KB, Lee K-H, Werner IM, 2009a. It's SHO time! short-sale price tests and market quality. *J. Finance* 64(1):37–73.

Diether KB, Lee K-H, Werner IM, 2009b. Short-sale strategies and return predictability. *Rev. Financial Stud.* 22(2):575–607.

Drechsler I, Moreira A, Savov A, 2020. Liquidity and Volatility. Working paper, Available at http://dx.doi.org/10.2139/ssrn.3133291.

Duflo E, Glennerster R, Kremer M, 2006. Using randomization in development economics research: A toolkit. *Handbook of Development Economics*, volume 4, 3895–3962.

Easley D, O'Hara M, Srinivas PS, 1998. Option volume and stock prices: Evidence on where informed traders trade. *J. Finance* 53(2):431–465.

Edwards AK, Reed AV, Saffi PAC, 2024. A survey of short selling regulations. *Rev. Asset Pricing Stud.* 14(4):613–639.

Eom KS, Binh KB, Kim J-S, 2011. Foreign investors' short sales and individual stock prices in the Korean stock market. *Korean J. Financial Management* 28(3):139–187, (in Korean).

Eom KS, Ok J, Park J-H, 2007. Pre-trade transparency and market quality. *J. Financial Markets* 10(4):319–341.

Eom KS, Seon J, Chang K-H, 2010. Relative efficiency of price discovery on an established new market and the main board: Evidence from Korea. *Asia-Pacific J. Financial Stud.* 39(4):459–494.

Ester M, Kriegel H-P, Sander J, Xu X, 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *Proc. 2^nd^ Int. Conf. on Knowledge Discovery and Data Mining*, 226–231.

Figlewski S, Webb GP, 1993. Options, short sales, and market completeness. *J. Finance* 48(2):761–777.

Friedman JH, 2001. Greedy function approximation: A gradient boosting machine. *Ann. Statist.* 29(5):1189–1232.

Goldstein I, Guembel A, 2008. Manipulation and the allocational role of prices. *Rev. Econom. Stud.* 75(1):133–164.

Grundy BD, Lim B, Verwijmeren P, 2012. Do option markets undo restrictions on short sales? Evidence from the 2008 short-sale ban. *J. Financial Econom.* 106(2):331–348.

Hamilton JD, 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57(2):357–384.

Hanley JA, McNeil BJ, 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143(1):29–36.

Hao X, Lee E, Piqueira N, 2013. Short sales and put options: Where is the bad news first traded? *J. Financial Markets* 16(2):308–330.

Hasbrouck J, 1993. Assessing the quality of a security market: A new approach to transaction-cost measurement. *Rev. Financial Stud.* 6:191–212.

Hasbrouck J, 1995. One security, many markets: Determining the contributions to price discovery. *J. Finance* 50(4):1175–1199.

Haslag P, Ringgenberg MC, 2023. The demise of the NYSE and Nasdaq: Market quality in the age of market fragmentation. *J. Financ. Quant. Anal.* 58(7):2753–2782.

Hendershott T, Jones CM, Menkveld AJ, 2011. Does algorithmic trading improve liquidity? *J. Finance* 66(1):1–33.

Huh S-W, Lin H, Mello AS, 2015. Options market makers' hedging and informed trading: Theory and evidence. *J. Financial Markets* 23(C):26–58.

Imbens GW, 2004. Nonparametric estimation of average treatment effects under exogeneity: A review. *Rev. Econom. Statist.* 86:4–29.

Jain C, Jain P, McInish TH, 2012. Short selling: The impact of SEC Rule 201 of 2010. *Financial Rev.* 47(1):37–64.

Jarrow R, 1980. Heterogeneous expectations, restrictions on short sales, and equilibrium asset prices. *J. Finance* 35(5):1105–1113.

Jiang GJ, Shimizu Y, Strong C, 2022. Back to the futures: When short selling is banned. *J. Financial Markets* 22(C):100735.

Jones CM, Lamont OA, 2002. Short-sale constraints and stock returns. *J. Financial Econom.* 66(2–3):207–239.

Kang J, Jeong G, 2018. Stock return predictability of the Amihud measure in the Korean stock market and trading volume. *Korean J. Financial Stud.* 47(4):543–577, (in Korean).

Khan MSR, 2024. Short-sale constraints and stock returns: A systematic review. *J. Cap. Markets Stud.* 8(1):43–66.

Lee BK, Lessler J, Stuart EA, 2010. Improving propensity score weighting using machine learning. *Statist. Medicine* 29:337–346.

LeLorier J, Gregoire G, Benhaddad A, Lapierre J, Derderian F, 1997. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *New Engl. J. Medicine* 337:536–542.

Levy B, 2022. Price Improvement and Payment for Order Flow: Evidence from A Randomized Controlled Trial. Working paper, Available at http://dx.doi.org/10.2139/ssrn.4189658.

Li F, Morgan KL, Zaslavsky AM, 2017. Balancing covariates via propensity score weighting. *J. Amer. Statist. Assoc.* 113(521):390–400.

Lo AW, MacKinlay AC, 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. *Rev. Financial Stud.* 1:41–66.

Locke PR, Sarkar A, 2001. Liquidity supply and volatility: Futures market evidence. *J. Futures Markets* 21(1):1–22.

Malinova K, Park A, 2024. Learning from DeFi: Would Automated Market Makers Improve Equity Trading? Working paper, Available at http://dx.doi.org/10.2139/ssrn.4531670.

Manaster S, Mann SC, 1996. Life in the pits: Competitive market making and inventory control. *Rev. Financial Stud.* 9(3):953–975.

Marsh IW, Payne R, 2012. Banning short sales and market quality: The UK's experience. *J. Banking Finance* 36(7):1975–1986.

McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF, 2013. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statist. Medicine* 32:3388–3414.

McCaffrey DF, Ridgeway G, Morral AR, 2004. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods* 9:403–425.

Miller EM, 1977. Risk, uncertainty, and divergence of opinion. *J. Finance* 32(4):1151–1178.

Morgan SL, Todd JL, 2008. A diagnostic routine for the detection of consequential heterogeneity of causal effects. *Sociol. Methodology* 38:231–281.

Moussawi R, Xu L, Zhou Z, 2024. A Market Maker of Two Markets: The Role of Options in ETF Arbitrage. Working paper, Available at https://ssrn.com/abstract=4395938.

Nagel S, 2012. Evaporating liquidity. *Rev. Financial Stud.* 25(7):2005–2039.

Pedregosa Fabian, Varoquaux Gaël, Gramfort Alexandre, Michel Vincent, Thirion Bertrand, Grisel Olivier, Blondel Mathieu, Prettenhofer Peter, Weiss Ron, Dubourg Vincent, Vanderplas Jake, Passos Alexandre, Cournapeau David, Brucher Matthieu, Perrot Matthieu, Duchesnay Édouard, 2011. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* 12:2825–2830.

Prado MP, Saffi PAC, Sturgess J, 2016. Ownership structure, limits to arbitrage, and stock returns: Evidence from equity lending markets. *Rev. Financial Stud.* 29(12):3211–3244.

Rosenbaum PR, Rubin DB, 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41–55.

Roth J, Sant'Anna PHC, Bilinski A, Poe J, 2023. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *J. Econometrics* 235(2):2218–2244.

Setoguchi S, Schneeweiss S, Brookhart MA, Glynn RJ, Cook EF, 2008. Evaluating uses of data mining techniques in propensity score estimation: A simulation study. *Pharmacoepidemiol. Drug Saf.* 17:546–555.

Shastri K, Thirumalai RS, Zutter C, 2008. Information revelation in the futures market: Evidence from single stock futures. *J. Futures Markets* 28(4):335–353.

Sikorskaya T, 2023. Institutional Investors, Securities Lending, and Short-selling Constraints. Working paper, Available at https://www.sikorskaya.net/research/.

Thomas LE, Li F, Pencina MJ, 2020. Overlap weighting: A propensity score method that mimics attributes of a randomized clinical trial. *J. Amer. Medical Assoc.* 323(23):2417–2418.

Venkataraman K, Waisburd AC, 2007. The value of the designated market maker. *J. Financial Quant. Anal.* 42(3):735–758.

Wang S-F, Lee K-H, 2015. Do foreign short-sellers predict stock returns? Evidence from daily short-selling in Korean stock market. *Pacific-Basin Finance J.* 32(C):56–75.

World Federation of Exchanges, 2022. WFE Derivatives Report 2021. April.

World Federation of Exchanges, 2024. FY 2023 Market Highlights Report. March.

Table A.1. Descriptive Statistics. This table presents the descriptive statistics for the variables used in the analysis, based on a sample of 6,441 observations. Samples are winsorized at the 1st and 99th percentiles to mitigate the effect of outliers. *Effective spread*, *price impact*, *execution shortfall*, and *pricing error variability* are scaled by 10,000. *Amihud illiquidity ratio* is scaled by one billion. *Execution shortfall* and *pricing error variability* are log-transformed. *Market basis gap* is standardized. Panel A shows market quality variables; Panel B shows variables used for propensity score estimation and panel regressions.

Category	Variable	Mean	Std	Min	25%	50%	75%	Max
Panel A. Market Quality Variables								
Liquidity	Effective Spread	8.4321	26.8207	-132.0606	4.7042	6.8122	10.1783	190.5658
	Price Impact	0.5143	2.5397	-11.6584	-0.253	0.4478	1.342	15.2304
	Amihud Illiquidity Ratio	12.9896	42.8591	0.0046	0.2921	1.0919	5.3843	333.8634
Volatility	1-Min Realized Volatility	0.6935	0.5347	0.0000	0.0005	1.0002	1.0005	2.4698
	5-Min Realized Volatility	0.6701	0.4974	0.0000	0.0004	1.0002	1.0004	2.0003
	10-Min Realized Volatility	0.6564	0.4782	0.0000	0.0003	1.0001	1.0003	1.2395
Price Efficiency	Variance Ratio	11.1887	72.8712	0.0021	0.6679	0.7582	0.8740	628.6234
	Execution Shortfall	4.4864	0.7932	2.5685	3.9320	4.4874	5.0519	6.2683
	Pricing Error Variability	1.0627	5.1154	-30.5905	1.0001	1.8209	2.6438	5.5324
Panel B. Variables for Propensity Scores and Panel Regression Analysis								
	Ask Market Depth	0.4456	0.2239	0.0763	0.2662	0.4149	0.5980	0.9798
	Bid Market Depth	0.4467	0.2057	0.0621	0.2894	0.4255	0.5918	0.9363
	Market Makers' Ask Trading Amount	0.2001	0.1667	0.0000	0.0727	0.1538	0.2857	0.7528
	Market Makers' Ask Trading Profit	304,484	882,748	-745,425	214	26,823	172,610	5,773,798
	Market Makers' Bid Trading Amount	0.1697	0.1482	0.0000	0.0565	0.1265	0.2487	0.6483
	Market Makers' Bid Trading Profit	34,672	733,712	-3,678,389	-29,041	2,242	63,674	3,564,424
	Market Basis Gap	-0.0324	0.9426	-11.0444	-0.1246	0.2214	0.3985	4.0214
	Max Price Change	0.1377	1.9598	-4.8478	-0.9221	0	0.9206	6.9565
	Trading Amount	18,542	44,979	27	815	3,178	14,981	315,070
	Trading Balance	-488	4,482	-26,325	-603	-42	253	16,046
	Sector	0.1705	0.3761	0	0	0	0	1
	Short-Selling Balance	0.0027	0.0049	0.0000	0.0003	0.0012	0.0029	0.0308
	Stock Trading Amount	605,907	1,692,665	0.0000	96,685	196,854	413,583	13,437,653
	Stock Volatility	-0.3419	2.0997	-5.1200	-1.3600	-0.5500	0.3100	11.2700

A. Appendix

This appendix provides supplementary materials supporting the main analysis. Figure A.1 justifies the selection of our sample period by conducting a volatility regime-switching analysis of daily returns of the KOSPI 200 stock index and the KOSPI 200 futures index using a Markov Regime Switching model. This analysis identifies periods of high and low volatility, demonstrating that the selected period falls within a stable low-volatility regime. Table A.1 presents summary statistics for key variables, offering a more detailed explanation of Table 1, including market quality measures and propensity-score calculation inputs. Table A.2 examines feature importance in propensity score estimation using a GBM model and Figure A.2 presents partial dependence plots for key features, specifically the *market basis gap* and *sector* variables, to illustrate their marginal effects on aggressive short selling.

To assess how data attributes affect classification accuracy, we analyze impurity-based importance, which measures a feature's contribution to reducing impurity when

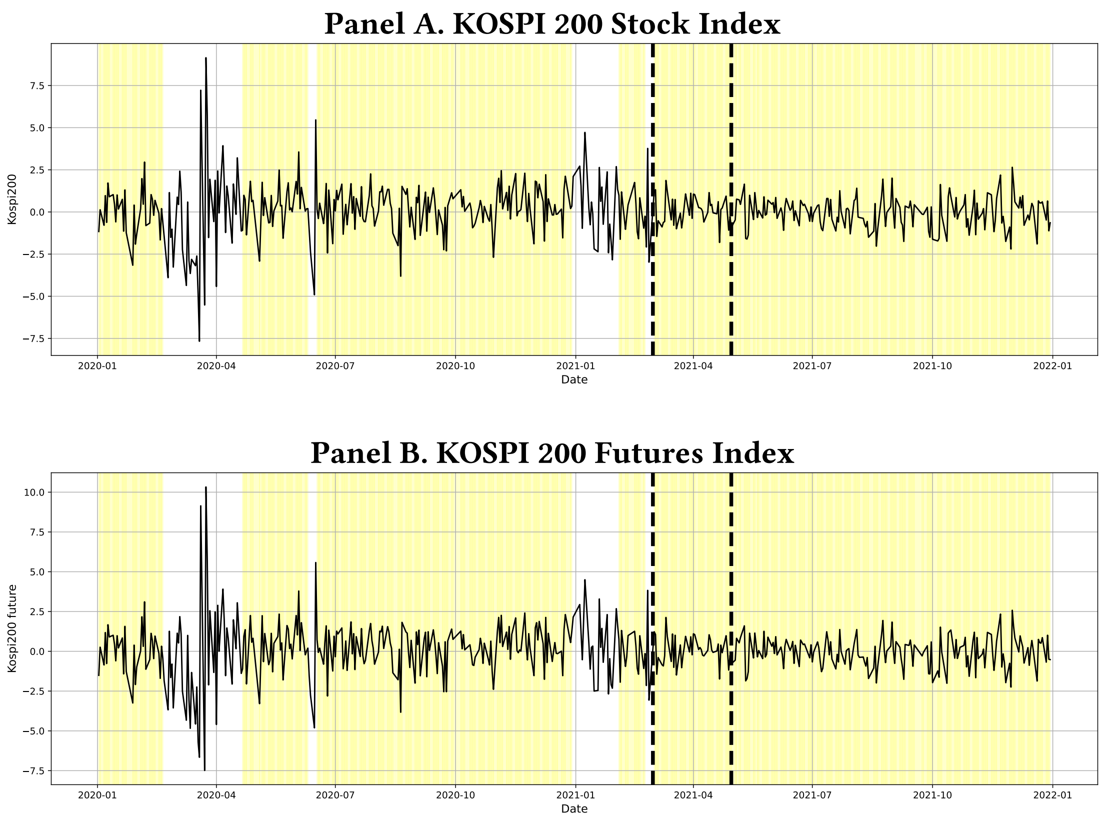

Figure A.1. Volatility Regime-Switching Analyses Panel A illustrates the results from the daily returns of the KOSPI 200 stock index, while Panel B presents the results from the daily returns of the KOSPI 200 futures index. Both panels use a Markov Regime Switching model with two regimes (high volatility and low volatility) to classify the returns (Hamilton, 1989). The low-volatility regime is shaded in yellow. The dotted vertical lines indicate the analysis period from March to April 2021; this period predominantly falls within the low-volatility regime of both return series, suggesting no significant market disruptions.

Table A.2. Feature Importance in Propensity Score Calculation. This table shows the relative importance of various features in predicting propensity scores for *Aggressive* and *Compliant* short-selling activities based on the GBM model estimated from the training dataset. All variables are log-transformed except for *market basis gap* and *sector*. 'S' denotes stock-related variables, while 'F' denotes futures-related variables. Importance Score represents each feature's percentage contribution to the model, based on its relative impact on impurity reduction. These scores are derived from the internal feature importance metrics provided by the fitted GBM model, reflecting how frequently and effectively each feature is used within the ensemble of decision trees. Higher scores indicate greater influence on model accuracy.

Features	**Importance Score**
Trading Amount (S)	19.99%
Market Basis Gap (S, F)	19.99%
Market Makers' Ask Trading Profit (F)	12.83%
Trading Amount (F)	9.88%
Ask Market Depth (F)	6.91%
Bid Market Depth (F)	5.48%
Sector (S, F)	4.91%
Market Makers' Bid Trading Amount (F)	4.39%
Account Weight (S)	4.33%
Market Makers' Ask Trading Amount (F)	3.20%
Max Price Change (F)	2.66%
Market Makers' Bid Trading Profit (F)	2.61%
Trading Balance (F)	1.67%
Volatility (S)	1.24%

used for a split (e.g., Gini impurity in classification or variance in regression); see Breiman et al. (2017) for reference. Specifically, this method quantifies the total impurity reduction attributed to a feature across all splits in the ensemble of decision trees. The resulting importance scores are then normalized to ensure that their sum equals one.

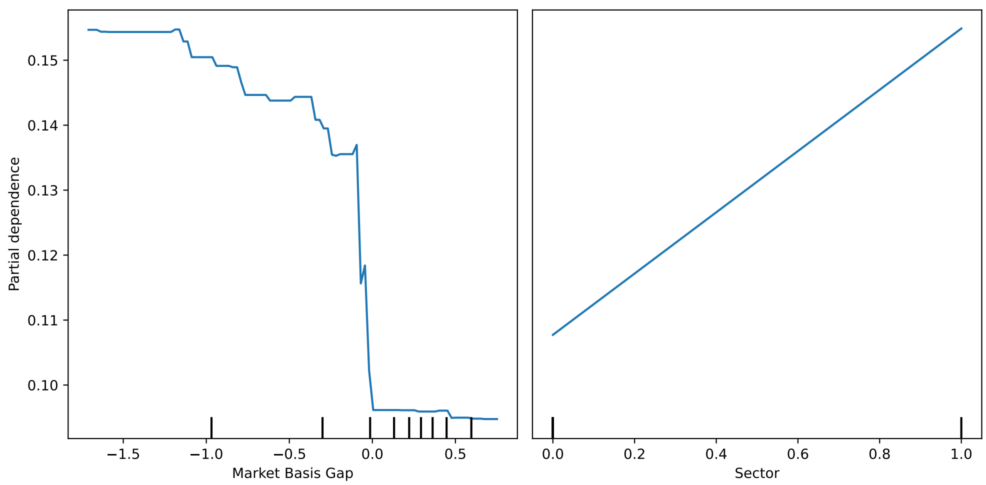

Figure A.2. Partial Dependence Plots for Important Features. This figure presents the Partial Dependence Plots (PDPs) for *market basis gap* and *sector*, which are key variables influencing the likelihood of aggressive short selling. PDPs illustrate the marginal effect of a feature on the predicted outcome of a model while averaging out the effects of all other features. The slope of each line indicates the direction and strength of the relationship between the feature and the model's predicted outcome: An upward slope reflects a positive association, a downward slope suggests a negative one, and a flat slope indicates no significant relationship. These plots provide interpretability by isolating the contribution of each individual variable to the model's behavior (Friedman, 2001).