Does High Frequency Market Manipulation Harm Market Quality?*

August 2023

Abstract: Manipulation of financial markets has long been a concern. With the automation of financial markets, the potential for high frequency market manipulation has arisen. Yet, such behavior is hidden within vast sums of order book data, making it difficult to define and to detect. We develop a tangible definition of one type of manipulation, spoofing. Using proprietary user-level identified order book data, we show the determinants of spoofing. Exploiting SEC Litigation Releases that exogenously reduce spoofing, we show causal evidence that spoofing increases return volatility, increases trading costs, and decreases price efficiency. The findings indicate that spoofing harms liquidity and price discovery.

JEL classification: G10, G12, G14

Keywords: high-frequency trading, market quality, market manipulation

1. Introduction

Modern financial markets are largely automated. With the increased automation, market participants can potentially distort markets to profitably induce short term price movements. One such high-frequency manipulation method is spoofing, which is defined as "bidding or offering with the intent to cancel the bid or offer before execution."[1] In September 2020, JPMorgan was fined $920 million for spoofing metals and U.S. Treasury futures, where it was suggested that spoofing is a common practice.[2,3] The frequency of spoofing activity in financial markets is an empirical question. In addition, the fact that spoofing should be unrelated to real information and therefore does not contribute to price discovery raises the question of how spoofing affects market quality. This paper quantifies the frequency of spoofing and tests whether it harms market quality.

Theory on the impact market manipulation should have on market quality is mixed. Skrzypacz and Williams (2021) address the determinants and market quality impacts of spoofing. They theoretically show that increased spoofing activity leads to slower price discovery, higher return volatility, and wider bid-ask spreads. A spoofing strategy impedes price discovery by driving prices away from fundamental values. Because deviations from fundamentals can be corrected, spoofing price movements induce reversals which then increase return volatility. At the same time, if spoofing drives prices away from fundamentals, adverse selection increases and market-makers are forced to raise spreads to remain profitable.

Some theoretical work argues against manipulation being feasible or that it can even improve market quality. Jarrow (1992) shows that when prices do not exhibit momentum,

[1] 2010 Dodd-Frank Act
[2] https://www.reuters.com/article/jp-morgan-spoofing-penalty-idINKBN26K325
[3] https://fortune.com/2022/07/20/former-jpmorgan-trader-reveals-how-his-mentor-taught-him-to-place-and-cancel-bogus-spoof-trades-manipulate-markets/

manipulation is not possible. Cherian and Jarrow (1995) show that a symmetric price response to manipulation renders it unprofitable. Other studies show that manipulation may be associated with improved market quality. Hanson and Oprea (2009) model a manipulator as a noise trader and show that the manipulation strategy encourages information acquisition as the profits to informed traders increase, thereby improving price accuracy. We empirically test these conflicting theories on the existence and effect of market manipulation.

We study Canadian equity markets using the proprietary IIROC dataset, which has trade and quote data with trader identification. We identify potential spoofing orders by applying six tractable filters to the data. We then examine the prevalence and determinants of spoofing in Canadian equity markets. We find that the median stock-day observation has 64 attempted spoofing orders, with 3 successful. We exploit variation in spoofing from SEC Litigation Releases to estimate the causal effect of spoofing on market quality. Our results are consistent with the theoretical predictions of Skrzypacz and Williams (2021). Spoofing leads to higher return volatility, higher transaction costs, and slower price discovery.

To discourage spoofing activity regulators strategically make the definition ambiguous. While is it not possible to perfectly identify spoofing orders, we draw from recent spoofing court cases[4] to develop a six-step filtering approach that identifies trade and order behavior consistent with spoofing. First, all spoofing orders are eventually deleted. Second, spoofing buy (sell) order prices must be greater (less) than or equal to one tick below (above) the prevailing NBB (NBO). We match potential spoofing orders to genuine orders, which are orders in the opposite direction from the same trader. Third, spoofing orders must be placed within one second of the genuine order. Fourth, the spoofing order volume must be higher than the genuine order volume. Fifth, the

[4] For example, *United States v. Coscia* and *United States v. Bases et al.*

spoofing orders must be cancelled within one second after genuine orders are executed or cancelled. Lastly, we require that during the second a spoofing order is placed, the trader does not actually trade in the same direction as the spoofing order. As it is very challenging to empirically distinguish market making from spoofing manipulation, we purposely use strict criteria that can distinguish between the two. A limitation of such a strict definition is that we likely undercount the true spoofing activity.

We begin the empirical analysis by documenting the prevalence and determinants of spoofing activity. Plotting spoofing against lagged market quality characteristics, we find that spoofing is most prevalent in stock-days with intermediate levels of transaction costs, intermediate levels of return volatility, and high levels of price efficiency.

Motivated by the theoretical predictions from Skrzypacz and Williams (2021), we next focus on the relationship between spoofing and market quality. We estimate OLS panel regressions of market quality measures on the attempted spoofing order volume scaled by trading volume, while controlling for lagged dollar spread, lagged price, lagged inverse price, absolute return, log of dollar volume, Amihud (2002) illiquidity, and stock and date fixed effects. Spoofing is positively associated with 1- and 5-minute return volatility, effective spreads, realized spreads, variance ratios, and the Hasbrouck (1993) pricing error. We also find that quoted spreads and adverse selection are negatively associated with spoofing activity.

There is a strong endogeneity problem. Spoofing traders likely endogenously select certain stocks and dates to spoof. For instance, Skrzypacz and Williams (2021) predict that spoofers endogenously choose to spoof when markets are not so illiquid that their spoofing orders can be identified by market makers but not so liquid that their spoofing orders are unable to move markets. We document a similar pattern. If spoofing activity is correlated with a stock's ex-ante liquidity, then our OLS estimates suffer from omitted variable bias, as ex-ante liquidity likely predicts market quality.

To overcome the endogeneity concern, we exploit SEC Litigation Releases as shocks to spoofing activity. We interpret market manipulation-related SEC Litigation Releases as positive shocks to the ex-ante legal risk of spoofing for stocks subject to SEC jurisdiction. In the three days after a release, spoofing in US cross-listed stocks decreases relative to stocks that are only listed on Canadian exchanges. Because SEC Litigation Releases predict spoofing activity but likely do not affect market quality directly, we instrument for spoofing with a difference in difference regression comparing the effect of SEC Litigation Releases on US cross-listed and Canada only stocks. The instrumental variables estimation shows that spoofing causes increased return volatility, increases variance ratios, and increases the Hasbrouck (1993) pricing error volatility. We also find weak evidence that spoofing raises quoted and effective bid-ask spreads, raises adverse selection, and lowers realized spreads.

We show that spoofing harms market quality at the intraday level. We regress 30-minute market quality measures on spoofing, lagged trading volume, lagged absolute return, and stock-day and 30-minute interval fixed effects. The strict fixed effects allow us to study the relation between spoofing and market quality *within* stock-day, which helps alleviate endogeneity concerns. The results indicate that spoofing strongly increases intraday return volatility, quoted spread, and effective spread. Spoofing decreases realized spreads and increases adverse selection, which suggests that spoofing activity harms liquidity providers. However, at high frequencies, spoofing has an economically insignificant effect on price efficiency when measured with the variance ratio.

We validate our spoofing measure by examining spoofing activity around the passage of the Dodd-Frank act. Namely, we observe a decrease in spoofing in US cross-listed stocks relative to stocks that are only listed on Canadian exchanges because of the more stringent anti-fraud provisions in Dodd-Frank that only apply to US cross-listed stocks. Because only US cross-listed stocks are subject to US regulations, Dodd-Frank should not affect spoofing in Canada-only stocks. The results suggest that increases in the ex-ante legal risk of spoofing can deter spoofing activity.

To alleviate concerns that our spoofing detection process captures legitimate orders and cancellations placed by market makers, we conduct a falsification test. We rely on key differences between spoofing and legitimate market making by HFTs. First, spoofing trading activity is one-sided, while market making trading is typically two-sided to provide liquidity. Second, spoofing strategies require that spoofing orders are cancelled quickly, while market makers place orders to maintain a two-sided market. For each stock-day, we measure market making activity as the proportion of orders from traders who have at least one outstanding order on each side of the limit order book at the end of each minute and place buy orders between 40% to 60% of the time. In OLS regressions of our market quality measures on market making activity, we find that market making activity is associated with improved market quality. This suggests that our spoofing measure does not capture legitimate market making.

Finally, we conduct a variety of robustness tests. We re-estimate our baseline IV results using alternative definitions of spoofing, such as the successful and failed spoofing order volume. Across the varying robustness checks the results remain economically consistent. We also re-estimate the IV results across different subsamples and exclude options settlement dates to address concerns that SEC litigation releases may also affect other types of manipulation.

This paper contributes to the extant literature on market manipulation (See Putnins, 2012 for a survey) and more specifically to the newer literature on high frequency market manipulation. There is a nascent theoretical literature on spoofing. In general, it is challenging to model limit order book dynamics (Parlour, 1998; Rosu, 2009). Theory has incorporated spoofing behavior into the equilibrium order book behavior. Skrzypacz and Williams (2021) provide an equilibrium model showing that spoofing behavior can harm liquidity, slow price discovery, and elevate volatility. Wang, Hoang, Vorobeychik, and Wellman (2021) also show that the presence of spoofers in an order book that is otherwise informative results in a decrease in investor welfare. Cartea, Jaimungal, and Wang (2020) model how spoofing can be used to increase an investor's revenue, and how potential legal fines can deter spoofing behavior. Using simulated limit order books, Withanawasam, Whigham, and Crack (2018) examine where manipulators may be more prevalent. Our study provides empirical tests of the theoretical implications of spoofing on market quality and confirms that spoofing harms market quality.

Legal scholars have argued more generally about the impact of spoofing. Fischel and Ross (1991) provide a framework for how the legal community analyzes manipulation in markets. They argue that it is difficult to identify manipulation without knowing trader intent. They propose that no trades should be considered manipulative, while behavior that gives a false sense of trading activity (i.e. wash trading or matched orders) is manipulative. McNamara (2016) tackles the ethical and legal implications of high frequency trading, which covers spoofing and other limit order based manipulation strategies. Miller and Shorter (2016) survey the literature on high frequency trading and market manipulation and discuss the regulatory and legislative reaction to crack down on behaviors such as spoofing. Canellos et al. (2016) provide an overview of spoofing cases that have occurred before and after Dodd-Frank. Fox, Glosten, and Guan (2021) provide a framework to consolidate the varying interpretations of what is and is not considered spoofing. Montgomery (2016) argues that spoofing may in fact improve the liquidity of financial markets. Dalko, Michael, and Wang (2020) argue that spoofing as a manipulative practice only arises because of behavioral biases of investors and microstructural systems.

The empirical work on spoofing is limited. The reason for the paucity of work on the topic is that it typically requires order book data with trader identifying information. That said, Tao, Day, Ling, and Drapeau (2022) have crafted a strategy to detect spoofing from public order books. Two other papers have identifying account information and study spoofing. Lee, Eom, and Park (2013) use data from Korea and show a positive correlation among spoofing and volatility and a negative correlation with market capitalization. Wang (2019) uses data from Taiwan futures and shows that spoofing is profitable and is correlated with higher volume, bid-ask spreads, and volatility. This paper makes two contributions to the empirical literature. First, we provide another tractable spoofing detection method that aims to be orthogonal to genuine market-making activity. Second, we are the first to provide causal evidence that spoofing negatively impacts market quality.

2. Data and Variable Construction

Our primary data source is the proprietary Investment Industry Regulatory Organization of Canada (IIROC) dataset. The data consists of trade and quote data for 127 Canadian stocks from May 3, 2010 to July 19, 2011. The sample is a volume stratified sample of Toronto Stock Exchange (TSX) stocks plus the TSX60 index constituents. Penny stocks and stocks with less than 20 active days are excluded. 46% of the firms in the sample are cross-listed in the US. We observe trades and quotes on the Toronto Stock Exchange. We also observe Alternative Trading System (ATS) activity through the Alpha (ALF), Chi-X (CHX), Omega (OMG), Pure (PTX), and MATCH Now (TCM) platforms.

The trade and quote data are timestamped at the 10-millisecond level and contain order submissions, amendments, cancellations, and executions. Importantly, trades and orders in the data have masked trader IDs that allow us to track individual trader positions and strategies across time. For each event, we observe trader ID, order ID, price, volume, NBB, NBO, exchange, and other information. Each order is assigned an order ID that can be used to track the status of an order over time. This is crucial for spoofing identification, as it allows us to track an individual trader's cancellations and amendments with precision. We require that each stock-day has at least $1 million in trading volume to remove very illiquid stocks. We drop observations with quoted spreads above 5% to remove potential data errors.[5]

2.1 Market Quality Measures

We construct liquidity and market quality measures from the IIROC data. We measure liquidity with time-weighed quoted spreads, volume-weighted effective spreads, volume-weighed realized spreads, volume-weighed adverse selection, and Amihud (2002) illiquidity. We measure volatility with 1- and 5-minute return volatility, and market quality with variance ratios and Hasbrouck (1993) pricing error σ.

We compute time-weighted quoted spreads by weighting $\frac{NBO-NBB}{NBBO\ midpoint}$ by the time each spread prevails for a given stock-day. We compute volume-weighted effective spreads by weighing $2 \times \frac{D_k(Price_k - NBBO\ midpoint_k)}{NBBO\ midpoint_k}$ by the volume at each trade, k, where D_k is a trade sign

[5] More details about the IIROC dataset can be found in the internet appendix for *The Competitive Landscape of High-Frequency Trading Firms* by Boehmer, Li, and Saar (2018).

indicator equal to 1 if the trade was buyer initiated and -1 if the trade was seller initiated. Because the data indicates the trade initiation direction, we sign trades directly. To approximate liquidity provision revenue, we compute volume-weighted realized spreads by weighing $2 \times \frac{D_k(Price_k - NBBO\ midpoint_{k,t+5})}{NBBO\ midpoint_k}$ by the volume at each at each trade, k, where $NBBO\ midpoint_{k,t+5}$ is the NBBO midpoint five minutes after trade k. Adverse selection is computed as the difference between the effective spread and realized spread. Amihud (2002) illiquidity is computed as the absolute value of daily returns divided by dollar volume for each stock day, multiplied by 10^6.

Return volatilities are computed at the 1- and 5-minute levels and are the standard deviation of returns using trading prices. We compute Lo and MacKinlay (1988) variance ratios with 1- and 30-minute return variances with $\left|1 - 30 \times \frac{Var_{1\ minute}(ret)}{Var_{30\ minute}(ret)}\right|$, a timing choice also used in Rösch, Subrahmanyam, and van Dijk (2016). We compute 1- and 30-minute returns with trade prices. Lastly, we compute the Hasbrouck (1993) pricing error σ. Similar to Boehmer and Kelley (2009), we estimate the VAR system with five lags and include four variables: log returns, trade sign indicator equal to 1 if the trade was buyer initiated and -1 if the trade was seller initiated, signed volume computed as the trade sign times the number of shares traded, and root signed volume computed as the trade sign times the square root of the number of shares traded. We set lagged variables to zero at the beginning of each day. Table 1 Panel A reports liquidity and market quality summary statistics at the stock-day level, while Panel B reports liquidity and market quality summary statistics at the 30-minute level.

INSERT TABLE 1 ABOUT HERE

2.2 Spoofing Measures

As the official definition of spoofing is likely strategically ambiguous, it is difficult to empirically measure the prevalence of spoofing activity. We draw our criteria from the following example of a trader who successfully executes a sell spoofing strategy: suppose a trader wants to buy shares of a stock. The NBB and NBO are currently $99 and $100, respectively. The trader wants to buy at a price less than $99 and will manipulate prices down. First, the trader places a buy order for the shares he wants to buy at $98.75, which is less than the prevailing NBO. He then rapidly places a high-volume limit sell order at a price lower than $100 (but higher than $99 to avoid immediate execution) to mimic selling pressure. The market responds to the false selling pressure by adjusting the NBB and NBO down. However, the trader immediately cancels the limit sell order before it can be executed. Because the market responds to the selling pressure, the NBB decreases and falls below $98.75, which results in the trader's buy order executing. Figure 1 describes this strategy graphically.

INSERT FIGURE 1 ABOUT HERE

Our example yields a more general definition. A trader who is spoofing the market will initially place a bona fide "genuine" buy limit order at a price lower than the current best bid price. After placing the genuine order, the trader will enter "spoofing" sell orders that will create the impression that the market is facing selling pressure. This will drive prices down and lead to the genuine order being executed. Finally, the spoofer will cancel the spoofing sell order. The same story holds with genuine sell orders and spoofing buy orders. We develop six filters to classify orders as potential spoofing orders.

We separately identify buy and sell spoofing orders. We also require that spoofing activity occurs during the trading hours of 9:30 AM to 4 PM. We describe the procedure for identifying spoofing buy orders in detail.[6] The spoofing identification procedure relies on visible trader IDs to track spoofing strategies.

We first search for spoofing orders without considering the other side's genuine orders. The first filter requires that spoofing orders are eventually deleted. As spoofing strategies consist of rapid entrance and cancellation of orders in the same direction, we expect that a spoofer will cancel a vast majority of their spoofing orders. Our spoofing detection strategy implicitly assumes that spoofing orders are not executed. Although it is likely that some spoofing orders are unintentionally executed, it is difficult to disentangle an executed spoofing order from a non-spoofing order. Second, if a spoofing order is to induce a market response, it must be somewhat aggressive. We require that buy spoofing order prices are greater than or equal to one tick below the previous NBB.

We match each potential buy spoofing order to potential sell (genuine) orders from the same trader ID.[7] Our third criteria requires that spoofing orders occur within one second after the genuine order is placed. This is consistent with a spoofing trader first entering a reasonable genuine order and then subsequently spoofing the market to induce a price response. For there to be a price

---

[6] The procedure to identify spoofing sell orders is nearly identical to the procedure used to identify buy orders. Switching "buy" with "sell" and changing the second filter to require that the spoofing sell order must be less than the NBO yields the spoofing sell order identification procedure.

[7] Note that our matching procedure can match multiple spoofing orders to a single genuine order. Our spoofing detection algorithm can therefore also capture layering activity, which regulators often use interchangeably with spoofing. Layering can be viewed as spoofing, but with multiple non-bona fide orders at different prices.

effect, spoofing orders again must be sufficiently aggressive. Our fourth filter captures this by requiring that each spoofing buy order's volume must be greater than the genuine order's volume. Spoofing occurs at high frequencies. Our fifth and most aggressive filter requires that spoofing orders are cancelled within one second after genuine orders are either cancelled or executed. Lastly, our sixth filter requires that for a given spoofing buy order, the trader ID must not have executed a buy order in the same second. This is consistent with the one-sided nature of spoofing. If a trader is trying to manipulate prices in one direction, it is unlikely that they will trade on their spoofing orders (and if they did, then the spoofing strategy would be much less profitable).

We define three types of spoofing: attempted, successful, and failed. Successful spoofing orders are spoofing orders with executed genuine orders, while failed spoofing orders have cancelled genuine orders. Attempted spoofing orders are the sum of the successful and failed spoofing orders. Our main measure of spoofing is the attempted spoofing order volume scaled by trading volume.

Table 1 Panel C presents the stock-day level summary statistics for spoofing activity. In our sample, the median stock-day has 64 attempted spoofing orders and 3 successful spoofing orders. Table 1 Panel D presents 30-minute level summary statistics for spoofing. The median number of attempted spoofs for a stock in a 30-minute interval is 3, while the median number of successful spoofs is 0. In untabulated results, we find that for a given stock-day, the median spoofer places three attempted spoofing orders. For a given trader-day, the median trader spoofs four stocks. High frequency traders place 57% of the average stock-day's spoofing orders.

Spoofing activity is right skewed, which suggests that spoofing may be heavily concentrated within certain time periods or stocks. We disaggregate successful and attempted spoofs into the buy and sell types and find that on average, selling spoofing activity is slightly more common than buying spoofing activity. This suggests that traders who wish to manipulate the market by spoofing tend to do so with downward price pressure.

2.3 Market-making Measure

A concern with our spoofing identification method is that we are measuring orders and cancellations associated with market making or liquidity provision activity. We generate a measure of liquidity provision to show that our results are likely not driven by market making. A trader-minute is considered market making if the proportion of buy orders is between 40% to 60% and the trader has at least one order outstanding at the end of the minute on each side of the market. Our market making measure is defined as the standardized percent of orders associated with market-making activity for each stock day.

2.4 Microstructure Controls

We compute average dollar spread, average price, and inverse price as microstructure controls in regression tests. Average price is computed as the dollar trading volume divided by share trading volume, and inverse price is equal to 1 divided by the average price. Average dollar spread is computed by multiplying the quoted spread by the average price.

3. Spoofing Activity

We begin by examining the determinants of spoofing activity graphically. We compute the average number of attempted spoofing orders for 15 lagged market quality quantiles. Skrzypacz and Williams (2021) predict that spoofing activity should be most active in markets with moderate liquidity. We measure liquidity with quoted spread, effective spread, realized spread, and adverse selection. We also show the relation between spoofing and lagged volatility and price efficiency measures. The results are shown in Figure 2.

INSERT FIGURE 2 ABOUT HERE

Panel A shows that spoofing tends to occur in stocks with lower ex-ante quoted spreads. However, spoofing is most prevalent in stock-days with intermediate ex-ante effective spreads, adverse selection, and realized spreads. This is consistent with the Skrzypacz and Williams (2021) prediction that spoofing should be the most prevalent in markets with intermediate levels of liquidity, as spoofers target sufficiently liquid markets to avoid being caught, while targeting sufficiently illiquid markets to be able to effectively influence prices.

Panel B presents results for ex-ante volatility. Spoofing occurs the most in stock-days with moderate levels of intraday return volatility. Spoofing in periods of low return volatility may lead to higher chances of being caught, while spoofing in periods of high return volatility is less likely to move prices in the desired direction. Panel C shows that spoofing occurs the most in stocks with lower inverse market quality levels. That is, spoofing is more prevalent when prices are more efficient. This is likely because spoofers target periods where their spoofing orders are more likely to be falsely impounded into prices as new information, such as when algorithmic trading is prevalent.

4. Relation between Spoofing and Market Quality

Guided by the theoretical predictions in Skrzypacz and Williams (2021), we examine the relation between spoofing activity and market quality. Namely, increased spoofing activity should be associated with higher return volatility, higher bid-ask spreads, and slower price discovery. We measure return volatility with 1 and 5-minute return volatility. We measure spreads with time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, and volume-weighted adverse selection. We measure price discovery with the variance ratio and Hasbrouck (1993) pricing error σ. All market quality variables except for the variance ratio are expressed in basis points. For each market quality measure, we estimate regressions of the following form:

$$Market\ Quality_{i,t} = \beta_1 Attempted\ Spoofing_{i,t} + \boldsymbol{\beta} X + \gamma_t + \zeta_i + \epsilon_{i,t}$$

Where $Attempted\ Spoofing_{i,t}$ is the standardized attempted spoofing order volume scaled by trading volume, and X is a vector of controls that includes lagged average dollar spread, lag average price, lag inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We denote date and stock fixed effects with γ_t and ζ_i, respectively.

We scale spoofing order volume by trading volume to more easily compare across stock-days with different levels of trading activity, as 500 shares of spoofing volume may have a different effect on market quality in a stock-day with 1 million vs 100 million shares of trading volume.

We include several microstructure and liquidity controls because the decision to spoof likely depends on a stock's ex-ante level of market quality (as shown in Figure 2). We include the lag of average price, inverse price, and average dollar spread. This is because spoofing may be easier to implement in stocks that are less tick constrained. However, we lag the microstructure variables to avoid controlling for a downstream affect, as spoofing may also directly affect price and dollar spread (Angrist and Pischke 2009).

Our controls for log dollar volume and Amihud (2002) illiquidity help control for contemporaneous liquidity, while the daily return control alleviates concerns that spoofing traders might tend to target stocks with high or low return magnitudes. Stock fixed effects sweep out time-invariant stock-specific variation, such as industry. Day fixed effects sweep out marketwide time variation, such as marketwide liquidity shocks.

INSERT TABLE 2 ABOUT HERE

The results in Table 2 show a positive relation between spoofing activity and most of the inverse market quality measures. Because the spoofing variable is standardized, the interpretation of β_1 is that a one standard deviation increase in scaled attempted spoofing orders is associated with a β_1 unit change in the dependent variable.

Spoofing increases return volatility. We find that a one standard-deviation increase in successful spoofing orders is associated with a 0.29 and 0.39 basis point increase in 1- and 5-minute return volatility, respectively. This is consistent with the idea that spoofing can move markets. If a spoofing trader can induce a temporary mispricing, then the process of inducing and correcting the manipulation will mechanically cause return volatility to increase.

Spoofing increases effective and realized bid-ask spreads but is associated with decreased quoted spreads and adverse selection. A one standard-deviation increase in successful spoofing orders is associated with a 0.1 basis point increase in the volume-weighted effective spread and

0.39 basis point increase in the volume-weighted realized spread. However, the spoofing coefficient on effective spreads is statistically insignificant. We find that spoofing is strongly negatively associated with quoted spreads: a one standard-deviation increase in successful spoofing orders is associated with a 0.16 basis point decrease in the quoted spread. This is likely because spoofing may occur inside the spread, therefore being NBBO improving and leading to temporary decreases in the quoted spread. Spoofing is negatively related to adverse selection: a one standard-deviation increase in spoofing is associated with a -0.27 basis point decrease in adverse selection.

Lastly, spoofing slows price discovery. A one standard-deviation increase in successful spoofing orders increases the variance ratio measure by 0.03 and Hasbrouck (1993) pricing error σ by 0.05 basis points. As the variance ratio measure increases, the ratio of 30 1-minute volatilities and 30-minute volatility deviates more from 1. This is evidence that increased spoofing activity drives price movements away from a random walk process, which suggests impeded price efficiency. The Hasbrouck (1993) procedure decomposes stock returns into random walk (efficient) and stationary (pricing error) components. Hasbrouck σ measures the variance of the pricing errors. Larger dispersion in pricing errors suggests a less efficient price process that tends to deviate more from true prices. Thus, the Hasbrouck σ result suggests that spoofing is also associated with lower price efficiency. However, the variance ratio and Hasbrouck σ estimates are economically small.

A potential shortcoming in our spoofing identification approach is that we cannot determine a trader's true intent and thus may be instead measuring genuine market-making activity. It is unlikely that genuine market-making activity will manifest in our measures because of our sixth filter: a trader must not place a spoofing order in the same second that they trade in that direction. Our sixth filter likely removes much market-making activity as market-making liquidity providers are more likely (or are required) to have balanced strategies. For example, the TSX appoints market makers who are required to maintain a two-sided market. Furthermore, if our spoofing variable measures market-making activity, then the results would contradict the existing literature on market-making. Market making should decrease spreads and improve market quality, which is the opposite of what we find. This suggests that our measure does not capture market-making activity. We provide further evidence that our results are not driven by market making with our analysis in Section 7.2.

Although we control for likely confounders and include stock and day fixed effects, it is possible that there are time-varying stock-specific unobservable or omitted variables that may bias our estimates. Thus, the results in this section can be viewed as associations between spoofing and market quality and are largely consistent with existing theoretical and empirical studies. Our finding that effective and realized spreads widen is consistent with Wang (2019), and the finding that return volatility is higher is consistent with Lee, Eom, and Park (2013). However, to our knowledge, we are the first to relate spoofing activity directly to price discovery measures such as variance ratios and Hasbrouck (1993) pricing errors.

5. Causal Effect of Spoofing on Market Quality

The results in Table 2 may suffer from omitted variable bias or reverse causality, as it is likely that spoofing traders endogenously respond to current liquidity or market quality conditions that may make spoofing strategies more profitable or effective. We exploit variation in spoofing induced by SEC litigation releases. The SEC issues litigation releases for its civil lawsuits in federal court.

The press releases range from initial charges filed by the SEC to final judgement announcements. We focus specifically on market manipulation related press releases that occur in the sample as shocks to spoofing activity.

SEC litigation releases likely affect the trading behavior of manipulative traders. We interpret litigation releases as positive shocks to the ex-ante legal risk of spoofing. Because regulators study limit order book data in market manipulation cases, a larger regulator presence increases the probability that manipulation is identified. If a spoofing trader observes that the SEC has begun or completed an investigation on market manipulation, the trader may infer heightened regulatory attention and thus a higher chance of being caught spoofing. The trader will thus reduce spoofing activity to reduce the chance of being caught.

We search the SEC Litigation Releases database for market manipulation releases.[8] A release is considered market manipulation if it contains the keyword "manipulation" and refers to stock price manipulation. For example, on September 24, 2010, the SEC charged four individuals with manipulating microcap stock prices. The traders allegedly engaged in a scheme to inflate two microcap stock prices and give a false sense of market liquidity in the stocks. Such events create a sense of heightened regulatory attention on market manipulation and should therefore discourage spoofing activity. We identify 22 SEC litigation releases on market manipulation in the sample period. To identify only the most severe shocks to the ex-ante legal risk of spoofing, we filter the list of releases to only include charges and final judgements. The final list consists of 10 SEC releases.

[8] https://www.sec.gov/litigation/litreleases.htm

Because we study the trading activity of cross-listed firms on Canadian exchanges, the analysis is only economically valid if SEC litigation releases can affect trading on Canadian markets. This is achieved through the Exchange Act of 1934's section on foreign securities exchanges.[9] Specifically, the provision on Foreign Securities Exchanges bans brokers and dealers from violating SEC regulations when trading on international exchanges if the stocks are "organized under the laws of" the United States. Because cross-listed stocks must comply with U.S. regulations, their stocks are likely protected from manipulation by U.S. and Canadian traders, even on Canadian exchanges. This is consistent with recent litigation. In *Harrington Global Opportunity Fund v CIBC World Markets Corporation,* U.S. and Canadian traders spoofed shares of Concordia International Corporation, a company cross listed in Canada (TSX) and the U.S. (NASDAQ), in 2016. The court acknowledged that a share of Concordia stock is the same whether it is traded on a U.S. or Canada exchange. Therefore, the court argued that it had jurisdiction over Canadian traders spoofing on Canadian exchanges because manipulating shares of Concordia would affect prices on NASDAQ.

We exploit the differential effect of SEC litigation releases on spoofing by comparing US cross-listed and Canada-only stocks. Because SEC litigation risk does not apply to Canada-only stocks, there should be a larger reduction in spoofing in US cross-listed stocks relative to Canada-only stocks. We use the differential effect of SEC litigation releases on spoofing in US cross-listed and Canada-only stocks to instrument for spoofing activity.

Our first stage estimate is the difference-in-differences regression of the standardized attempted spoofing order volume scaled by trading volume on the interaction between $US\ Listed_i$, which is an indicator equal to 1 if the stock is cross-listed in the US, and $Litigation_t$, which is an

[9] 15 U.S. Code § 78dd

indicator equal to 1 if day t is one to three days after a SEC litigation release on market manipulation. We choose a short period for $Litigation_t$ to avoid capturing slower moving reductions in manipulation which may plausibly improve market quality, such as insider trading or corporate misconduct. We include controls for lagged average dollar spread, lag average price, lag inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We also include stock and date fixed effects and cluster standard errors by stock. The first-stage results are presented in Table 3. The instrument is valid if it satisfies both the relevance and exclusion restrictions.

INSERT TABLE 3 ABOUT HERE

The first-stage results in Table 3 show that the instrument is powerful. The coefficient on $US\ Listed_i \times Litigation_t$ shows that in the three days after a SEC litigation release, US cross-listed stocks experience a 0.18 standard deviation decline in spoofing relative to Canada-only stocks. This is consistent with the hypothesis that SEC litigation releases cause spoofing activity to decrease in US cross-listed stocks, as traders reduce their spoofing activity in response to heightened legal risk. The T-statistic on $SEC_t \times Treat_i$ is -5.71 and the Kleibergen-Paap rk Wald F statistic (shown in Table 4) is greater than 32. The highly significant coefficient on the instrument and large Kleibergen-Paap rk Wald F statistic suggest that the relevance condition is satisfied. Figure 3 shows this relation graphically.

INSERT FIGURE 3 ABOUT HERE

The exclusion restriction requires that $US\ Listed_i \times Litigation_t$ only affects market quality through spoofing. Threats to exclusion would have to be correlated with both $US\ Listed_i \times Litigation_t$ and market quality and orthogonal to the second stage controls. While it cannot be empirically tested, it is challenging to think of alternative possible channels by which SEC litigation releases affect market quality other than through lowering market manipulation activity. One potential concern is that the IV affects high frequency market manipulation other than spoofing, such as short selling manipulation, settlement manipulation, and wash trading. To alleviate these concerns, we conduct robustness tests that suggest that the results are not driven by short selling manipulation or settlement manipulation. Furthermore, other types of manipulation such as wash trading create a false impression of liquidity. Therefore, if we observe that spoofing harms market quality, this would be despite any decreases in wash trading which may have otherwise improved short term market quality.

The second stage estimates are shown in Table 4. We regress the market quality measures from Table 2 on the predicted standardized spoofing values from the first stage estimate in Table 4. We again control for lagged average dollar spread, lag average price, lag inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We also include stock and date fixed effects and cluster standard errors by stock.

INSERT TABLE 4 ABOUT HERE

The results show that instrumented spoofing has a positive relation with return volatility. A one standard-deviation increase in spoofing causes a 2.5 basis point and 3.41 basis point increase in 1-minute and 5-minute return volatility, respectively. Spoofing has a statistically weak positive relation with quoted spread, effective spread, and adverse selection. This is consistent with spoofing leading to higher transaction costs. However, spoofing has a statistically weak negative relation with realized spreads, which is consistent with spoofing decreasing the profits of liquidity providers, since spoofing has typically targeted high frequency traders who may act as de-facto market makers. Spoofing causes higher variance ratios and Hasbrouck σ. A one standard-deviation increase in spoofing leads to a 0.99 increase in the variance ratio and a 1.36 basis point increase in Hasbrouck σ, which is evidence that spoofing harms price discovery. The results suggest that spoofing harms market quality.

6. Intraday Spoofing and Market Quality

We turn to the intraday relation between spoofing and market quality. In daily-level tests, we document an economically strong relation between spoofing and volatility and price discovery measures. However, it is likely that spoofing has a stronger intraday effect given its high frequency and fast time to completion. We measure spoofing at 30-minute intervals as the sum of attempted spoofing order volume scaled by daily trading volume. We again standardize the spoofing measure for ease of interpretation. We measure market quality with 1 and 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted adverse selection, and the variance ratio.

To estimate the intraday relation between spoofing and market quality, we estimate the following regression equation for each market quality measure:

$$Market\ Quality_{i,t,j} = \beta_1 Attempted\ Spoofing_{i,t,j} + \boldsymbol{\beta} X + \theta_{it} + \phi_j + \epsilon_{i,t,j}$$

Where $Attempted\ Spoofing_{i,t,j}$ is the standardized attempted spoofing order volume scaled by daily trading volume for stock i on day t during 30-minute interval j, and X is a vector of controls that includes lagged volume and absolute return. We include stock-day fixed effects with θ_{it} and 30-minute interval fixed effects with ϕ_j, respectively. Results are presented in Table 5.

INSERT TABLE 5 ABOUT HERE

Because there are multiple observations for each stock-day, we are able to apply an aggressive set of controls to help mitigate omitted variable bias. We first include the lag of trading volume and absolute return to alleviate concerns that spoofers may target periods with different levels of trading volume or return magnitudes. Importantly, the stock-day fixed effects sweep out any time varying (at the daily level) and firm specific characteristics, while the 30-minute interval fixed effects mitigate concerns that spoofing and market quality may be higher or lower during different times of day (Lee, Eom, and Park 2013). The aggressive fixed effects allow us to examine the effects of market quality within stock-day, which helps alleviate endogeneity concerns. However, we are unable to use the SEC Litigation Release IV because stock-day fixed effects are perfectly multicollinear with the instrument.

Consistent with the daily results, intraday spoofing leads to significant increases in return volatility. A one standard-deviation increase in 30-minute spoofing leads to a 1.09 basis point increase in 1-minute volatility and a 2.26 basis point increase in 5-minute volatility. Spoofing also significantly increases spreads. A one standard-deviation increase in 30-minute spoofing is associated with a 0.10 basis point increase in quoted spread and a 0.25 basis point increase in effective spread. This is evidence that spoofing increases transaction costs. Again consistent with the idea that spoofing targets liquidity providers, a one standard-deviation increase in spoofing significantly reduces realized spreads by 0.56 basis points and increases adverse selection by 0.84 basis points. Finally, intraday spoofing leads to a negative but economically insignificant change in the variance ratio. The results suggest that spoofing harms market quality at the intraday level.

7. Robustness

We apply a battery of robustness tests to validate our spoofing measures, ensure that our results are not driven by market-making activity or our choice of spoofing measure, and support the exclusion restriction assumption in the IV analysis.

7.1 Dodd-Frank and Spoofing

We explore the relation between spoofing and regulation by studying spoofing variation induced by the Dodd-Frank act. Because our sample is comprised of Canadian stocks, some of which are cross listed in the U.S., we interpret Dodd-Frank as a shock to the U.S. cross listed stocks' legal risk of spoofing. We use a difference-in-differences framework to show that anti-spoofing regulations can deter spoofing activity.

The 2010 Dodd-Frank Act was enacted on July 21, 2010 in the aftermath of the Global Financial Crisis. The legislation provided broad reforms to the US financial industry and was primarily related to regulating banks and mortgage markets. However, Dodd-Frank also increased investor protection in financial markets. The act's amendment to the Commodity Exchange Act was the first legislation to explicitly ban spoofing activities, although it was directed at commodity futures markets. Dodd-Frank also strengthened the antifraud provisions of the Securities Exchange

Act of 1934. Section 929L's amendment to §15(c)(1)(A) extended the ban on broker or dealer manipulation from off-exchange markets to brokers or dealers operating on national securities exchanges.

There are two possible channels by which Dodd-Frank affects Canada-only stocks compared to cross-listed stocks. First, there is a direct effect through the amendment to §15(c)(1)(A) that more clearly banned on-exchange manipulation by brokers and dealers. Second, there is an attention effect from the spoofing provision in Dodd-Frank, which was the first regulation to formally discuss (and ban) spoofing. While the regulation explicitly banned spoofing in commodity futures markets and allowed enforcement by the CFTC, it is plausible that U.S. traders decreased their spoofing activity in cross-listed stocks relative to Canada-only stocks due to expected heightened regulatory attention from U.S. regulators. Because manipulators likely invest significant resources in minimizing legal risk, it is plausible that most manipulators were aware of the manipulation-related rule changes. Furthermore, around the time of the act's passage, several law firms improved accessibility by issuing condensed summaries of the rule changes.

To avoid confounders due to a long time-horizon, we restrict the sample to the first 100 days of the sample, with July 21, 2010, the day Dodd-Frank was signed into law, being the 48th day in the sample. The results are robust to both shorter and longer windows. We exclude stocks with financial and insurance sector NAICS codes to alleviate concerns that Dodd-Frank may have affected market quality in a way that is correlated with our spoofing measure.

Because Dodd-Frank was anticipated by market participants, the choice of the "post" period in the difference-in-differences analysis is ex-ante ambiguous. In the sample period, there are three events that may also lead to a downward shift in spoofing in cross-listed stocks. On May 20, 2010, Dodd-Frank was passed in the senate with a 50 to 39 vote. On June 10, 2010, the conference report was filed for discussion and the act moved to the conference committee stage. Lastly, Dodd-Frank was signed into law on July 21, 2010. We estimate the average level of spoofing in US cross-listed stocks and Canada-only stocks in Figure 4. We define spoofing as the sum of attempted spoofing order volume scaled by trading volume.

INSERT FIGURE 4 ABOUT HERE

From the start of the sample to May 20, spoofing in US cross-listed and Canada-only stocks is flat across time. Following the passage of Dodd-Frank in the senate on May 20, there is a temporary decline and subsequent increase in spoofing in US cross-listed stocks. After Dodd-Frank entered the conference committee stage on June 10, spoofing declines and then remains near the pre-committee levels. However, once Dodd-Frank is signed into law, there is a noticeable decline in spoofing activity. Because spoofing starts to have a noticeable change (in US cross-listed relative to Canada-only stocks) after Dodd-Frank is signed into law, we use July 21, 2010 to define the start of the post period in the difference-in-differences analysis.

We verify formally that spoofing declines in Table 6. We measure spoofing in four ways: the count of attempted spoofing orders, count of successful spoofing orders, attempted spoofing order volume, and successful spoofing order volume. We regress the four spoofing measures on a $US\ Listed_i \times Post_t$ variable that is equal to 1 if stock i is cross-listed in the US and if day t is on or after July 21, 2010. The results show that there is a strong negative effect of Dodd-Frank on US cross-listed stocks relative to Canada-only stocks. The number of attempted spoofing orders falls by 2076.81, while the number of successful spoofing orders falls by 20.07. The scaled results also show that spoofing declines. These results are consistent with the idea that the anti-spoofing and anti-manipulation provisions in Dodd-Frank deterred potential manipulators from manipulating US cross-listed stocks, while Canada-only stocks were unaffected.

INSERT TABLE 6 ABOUT HERE

These results suggest that the anti-spoofing and anti-manipulation provisions in Dodd-Frank deterred potential manipulators from manipulating US cross-listed stocks, while Canada-only stocks were relatively unaffected. We interpret Dodd-Frank as a positive shock to the ex-ante legal risk of spoofing for U.S. cross-listed stocks. The results also provide support for the validity of our measure. If the true level of spoofing falls because of Dodd-Frank, then a valid proxy for the true level of spoofing should also fall.

7.2 Market Making

One potential concern is that the spoofing detection filters pick up bona fide market making activity. We conduct a falsification test to show that unlike spoofing, market-making activity improves market quality.

We rely on key differences between spoofing and legitimate market making by HFTs. First, spoofing trading activity is one-sided, while market making trading is typically two-sided to provide liquidity. Second, spoofing strategies require that spoofing orders are cancelled quickly, while market makers place orders to maintain a two-sided market. For each stock-day, we measure market making activity as the proportion of orders from traders who have at least one outstanding order on each side of the limit order book at the end of each minute and place buy orders between 40% to 60% of the time.

We repeat the OLS estimations from Table 2 with market-making activity instead of spoofing. The market-making variable is defined as the standardized percentage of orders associated with market-making activity (as defined in Section 2.3).

INSERT TABLE 7 ABOUT HERE

Table 7 shows that market-making activity decreases return volatility, lowers spreads, and lowers variance ratios and Hasbrouck σ. These results are generally consistent with the existing literature that increased algorithmic trading improves liquidity and price discovery (Hendershott, Jones, and Menkveld, 2011; Brogaard, Hendershott, and Riordan, 2014). However, the relation between market making and adverse selection is positive. This may be because in equilibrium, informed traders endogenously select to trade when there is the most liquidity provision available. The marking making results suggest that the spoofing measures are likely not capturing genuine market-making activity.

7.3 Alternative Spoofing Definitions

The main results measure spoofing as the standardized attempted spoofing volume scaled by trading volume. We show that the results are robust to alternate definitions of spoofing at the daily and intraday levels. Namely, we turn to successful and failed spoofs.

We define successful spoofing orders as attempted spoofing orders that also result in the genuine side being executed. We define failed spoofing orders as attempted spoofing orders that result in the genuine side being cancelled. We again scale each measure by the daily trading volume and standardize the measure for interpretation. For each alternative spoofing measure, we re-estimate the daily SEC litigation release IV approach and intraday regression approach.

The SEC litigation IV results are robust across the two different spoofing measures. Table 8 presents first stage estimates of the SEC litigation IV across the two alternate measures. The first stage coefficients on $US\ Listed_i \times Litigation_t$ show that SEC litigation releases cause spoofing to fall in US cross-listed stocks relative to Canada-only stocks. The Kleibergen-Paap rk Wald F statistic (shown in Table 9) is greater than 6 for successful spoofs and is greater than 24 for failed spoofs. This indicates that the first stage is weaker for successful spoofing activity.

INSERT TABLE 8 ABOUT HERE

Table 9 presents the second stage estimates for the SEC litigation IV. Panel A presents results for successful spoofing, while Panel B presents results for failed spoofing. The successful spoofing results in Panel A are statistically weaker than in Table 4, which is likely a result of the lower variation in successful spoofing activity. Although statistically weak, spoofing tends to increase return volatility, variance ratio, and Hasbrouck σ. Further, spoofing increases quoted spread, effective spread, and adverse selection, while decreasing realized spreads. The failed spoofing results in Panel B show a similar pattern. The coefficients in Panel A are all larger in magnitude than the corresponding coefficients in Panel B, which suggests that successful spoofing activity may have a larger affect on market quality relative to failed spoofing.

INSERT TABLE 9 ABOUT HERE

We also repeat the intraday analysis with successful and failed spoofing. We reestimate the regression specification from Table 5 but replace the attempted spoofing measure with either the scaled order volume from successful spoofing or the scaled order volume from failed spoofing. Table 10 presents the results.

INSERT TABLE 10 ABOUT HERE

Table 10 Panel A presents results for successful spoofing. The results provide strong evidence that spoofing is associated with higher volatility, quoted spread, effective spread, and adverse selection. Again, spoofing is associated with significantly lower realized spreads. Panel B presents the results for attempted spoofing and yields similar conclusions. However, the magnitudes on successful spoofing are larger, except for the quoted spread and variance ratio. This again indicates that successful spoofing activity may have a larger adverse effect on market quality.

7.4 Exclusion Restriction Robustness

The exclusion restriction requires that our instrument only affects market quality through spoofing activity. Because there are other types of market manipulation that may be affected by SEC litigation releases and correlated with spoofing activity, we make two modifications to our IV specification to mitigate this concern.

First, SEC litigation releases may lead to a decrease in short selling manipulation. If short selling manipulation is correlated with our spoofing measure, then the results may be biased.

Because sell spoofs may be correlated with short selling manipulation, the results may be contaminated with changes in short selling manipulation. We therefore re-estimate the IV specification results using buy spoofs only and sell spoofs only and find that the results are economically consistent with the estimates in Table 4. We also split the sample on above and below median lagged Amihud (2002) illiquidity. The IV results are economically similar in each of the two subsamples, although statistically weaker. Because it is likely more difficult to manipulate with short selling in illiquid stock-days, the concern of short selling manipulation is lessened in the illiquid subsample.

Second, SEC litigation releases may lead to a decrease on settlement manipulation around options expirations dates. Because Canadian equity options expire on the third Friday of each month, traders may manipulate spot prices to profitably trade options. We remove the third Fridays of each month and find that our results are again unchanged.

8. **Conclusion**

We document evidence of spoofing behavior in Canadian equity markets and provide causal evidence that spoofing harms market quality. Consistent with the theoretical predictions in Skrzypacz and Williams (2021), spoofing increases return volatility, increases transaction costs, and slows price discovery.

We develop a tractable six-step filtering process to identify spoofing orders and study the prevalence of spoofing. Consistent with Skrzypacz and Williams (2021), we show that spoofing activity is single-peaked in liquidity when measured with spreads and volatility.

OLS regressions show that on average, spoofing activity is associated with worse market quality. Using SEC Litigation Releases, we exploit the variation in spoofing in US-Canada cross-listed and Canada-only stocks in an instrumental variables framework to provide causal evidence that spoofing harms market quality. We estimate the relation between spoofing and market quality within stock-day and show similar results.

This paper makes two contributions to the literature. First, we provide another tractable spoofing detection method that aims to be orthogonal to genuine market-making activity. Second, motivated by the theoretical predictions in Skrzypacz and Williams (2021), we are the first to provide causal evidence that spoofing harms market quality.

References

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

Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-frequency trading and price discovery. *Review of Financial Studies*, 27(8), 2267-2306.

Canellos, G. S., Rangwala, T. S., Walfish, D. R., Jou, J. K., & Palladino, S. L. (2016). The law surrounding spoofing in the derivatives and securities markets. In *FIA L&C Conference*.

Cartea, Á., Jaimungal, S., & Wang, Y. (2020). Spoofing and price manipulation in order-driven markets. *Applied Mathematical Finance*, 27(1-2), 67-98.

Cherian, J. A., & Jarrow, R. A. (1995). Market manipulation. *Handbooks in Operations Research and Management Science*, 9, 611-630.

Dalko, V., Michael, B., & Wang, M. (2020). Spoofing: effective market power building through perception alignment. *Studies in Economics and Finance*, 37(3), 497-511.

Fischel, D. R., & Ross, D. J. (1991). Should the law prohibit manipulation in financial markets. *Harvard Law Review*, 105, 503.

Fox, M. B., Glosten, L. R., & Guan, S. S. (2021). Spoofing and its Regulation. *Columbia Business Law Review*, Forthcoming.

Hanson, R., & Oprea, R. (2009). A manipulator can aid prediction market accuracy. *Economica*, 76(302), 304-314.

Hasbrouck, J. (1993). Assessing the quality of a security market: A new approach to transaction-cost measurement. *Review of Financial Studies*, 6(1), 191-212.

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity?. *Journal of Finance*, 66(1), 1-33.

Jarrow, R. A. (1992). Market manipulation, bubbles, corners, and short squeezes. *Journal of Financial and Quantitative Analysis*, 27(3), 311-336.

Lee, C. M., Mucklow, B., & Ready, M. J. (1993). Spreads, depths, and the impact of earnings information: An intraday analysis. *Review of Financial Studies*, 6(2), 345-374.

Lee, E. J., Eom, K. S., & Park, K. S. (2013). Microstructure-based manipulation: Strategic behavior and performance of spoofing traders. *Journal of Financial Markets*, 16(2), 227-252.

McNamara, S. (2016). The law and ethics of high-frequency trading. *Minnesota Journal of Law Science and Technology*, 17, 71.

Miller, R. S., & Shorter, G. (2016). High frequency trading: Overview of recent developments (Vol. 4). *Washington, DC: Congressional Research Service*.

Montgomery, J. D. (2016). Spoofing, market manipulation, and the limit-order book. *Working Paper*.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1), 41-66.

Parlour, C. A. (1998). Price dynamics in limit order markets. *Review of Financial Studies*, 11(4), 789-816.

Putniņš, T. J. (2012). Market manipulation: A survey. *Journal of Economic Surveys*, 26(5), 952-967.

Rösch, D. M., Subrahmanyam, A., & Van Dijk, M. A. (2017). The dynamics of market efficiency. *Review of Financial Studies*, 30(4), 1151-1187.

Roşu, I. (2009). A dynamic model of the limit order book. *Review of Financial Studies*, 22(11), 4601-4641.

Tao, X., Day, A., Ling, L., & Drapeau, S. (2022). On detecting spoofing strategies in high-frequency trading. *Quantitative Finance*, 22(8), 1405-1425.

Wang, X., Hoang, C., Vorobeychik, Y., & Wellman, M. P. (2021). Spoofing the limit order book: A strategic agent-based analysis. *Games*, 12(2), 46.

Wang, Y, (2019). Strategic Spoofing Order Trading by Different Types of Investors in Taiwan Index Futures Market. *Journal of Financial Studies*, 27(1), 65- 104

Williams, B., & Skrzypacz, A. (2021). Spoofing in Equilibrium. *Working Paper.*

Withanawasam, R., Whigham, P., & Crack, T. F. (2018). Are Liquid or Illiquid Stocks More Easily Manipulated? The Impact of Manipulator Aggressiveness. *Working Paper*.

Appendix

Table A1. Variable definitions

Variable	Definition
Spoofing measures	
Attempted spoofs	Order volume from attempted spoofing orders as defined by the procedure in Section 2.2, scaled by daily trading volume. Includes both successful and unsuccessful spoofs, meaning that the associated genuine order does not have to be executed.
Successful spoofs	Order volume from successful spoofing orders as defined by the procedure in Section 2.2, scaled by daily trading volume. Includes only successful spoofs, meaning that the associated genuine order must be executed.
Failed spoofs	Order volume from failed spoofing orders as defined by the procedure in Section 2.2, scaled by daily trading volume. Includes only unsuccessful spoofs, meaning that the associated genuine order must be cancelled.
Market characteristics	
1-minute return volatility	Standard deviation of 1-minute returns.
5-minute return volatility	Standard deviation of 5-minute returns.
Quoted spread	Time-weighted quoted spread, where each quoted spread is $\frac{NBO-NBB}{NBBO\ midpoint}$.
Effective spread	Volume-weighted effective spread, where each effective spread is $2 \times \frac{D_k(Price_k - NBBO\ midpoint_k)}{NBBO\ midpoint}$. D_k is a trade sign indicator equal to 1 if the trade was buyer-initiated, and -1 if the trade was seller-initiated.
Realized spread	Volume-weighted realized spread, where each realized spread is $2 \times \frac{D_k(Price_k - NBBO\ midpoint_{k,t+5})}{NBBO\ midpoint_k}$. D_k is a trade sign indicator equal to 1 if the trade was buyer-initiated, and -1 if the trade was seller-initiated. $NBBO\ midpoint_{k,t+5}$ is the NBBO midpoint five minutes after trade k occurs.
Variance ratio	Lo and MacKinlay (1988) variance ratios using 1 and 30-minute return variances: $\left|1 - 30 \times \frac{Var_{1\ minute}(ret)}{Var_{30\ minute}(ret)}\right|$.
Hasbrouck σ	Standard deviation of pricing errors from VAR system with five lags and four variables: log returns, trade sign indicator equal to 1 (-1) if the trading price is buyer (seller) initiated, signed volume computed as the trade sign times the number of shares traded, and root signed volume computed as the trade sign times the square root of the number of shares traded.
Dollar trading volume	Total trading volume.
Absolute return	Absolute value of return for the trading day.

Market-making	Percent of orders associated with market-making activity. As defined in Section 2.3, market-making trader-minutes must have proportion of buy orders between 40% to 60% and must have an outstanding order at the end of the minute for each side.
Microstructure Controls	
Average price	Dollar trading volume divided by share trading volume.
Inverse price	1 / Average price
Dollar spread	Average price × quoted spread

Figure 1: Spoofing Example

Figure 1 provides a graphical representation of the sell spoofing example described in Section 2.2.

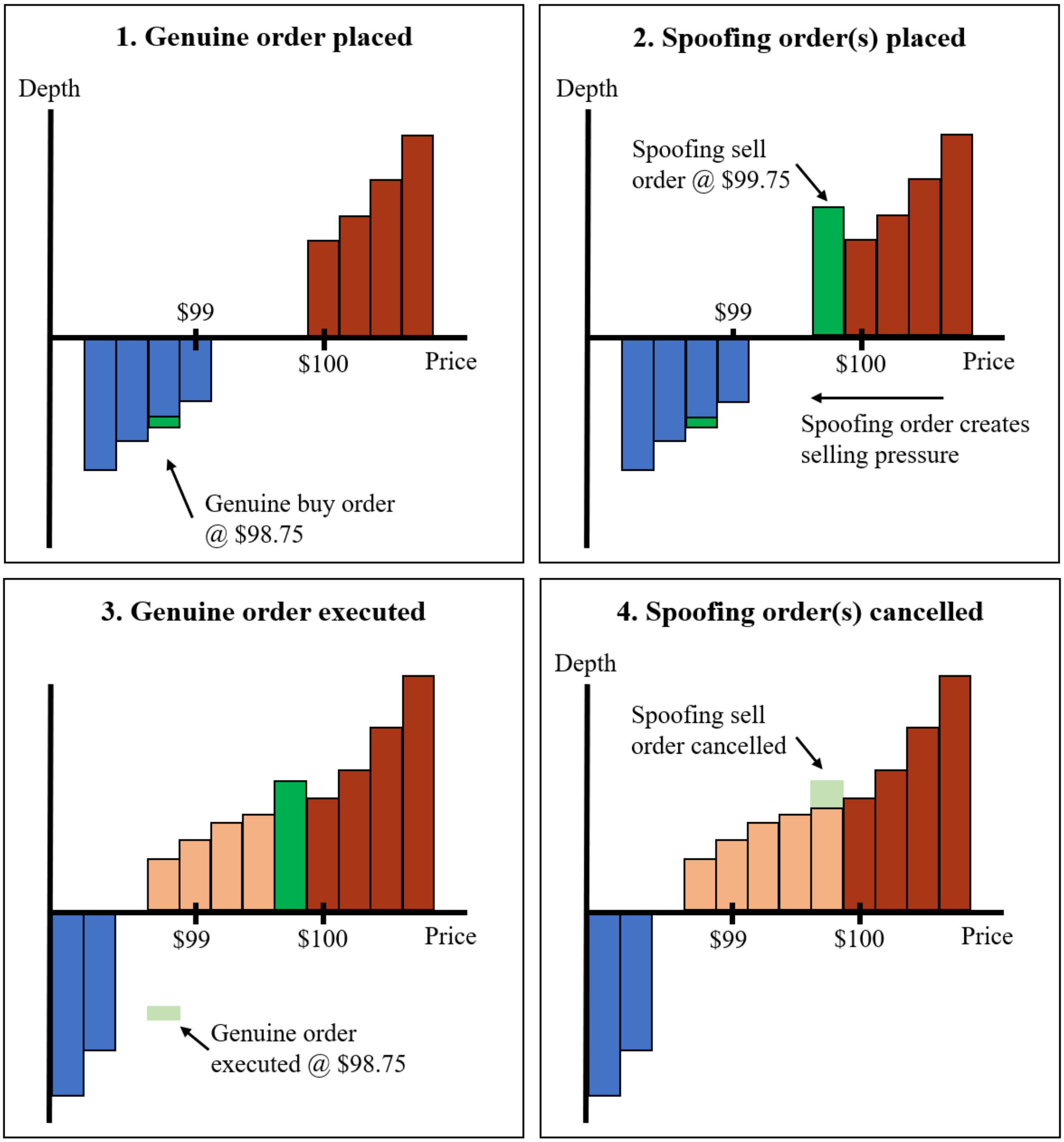
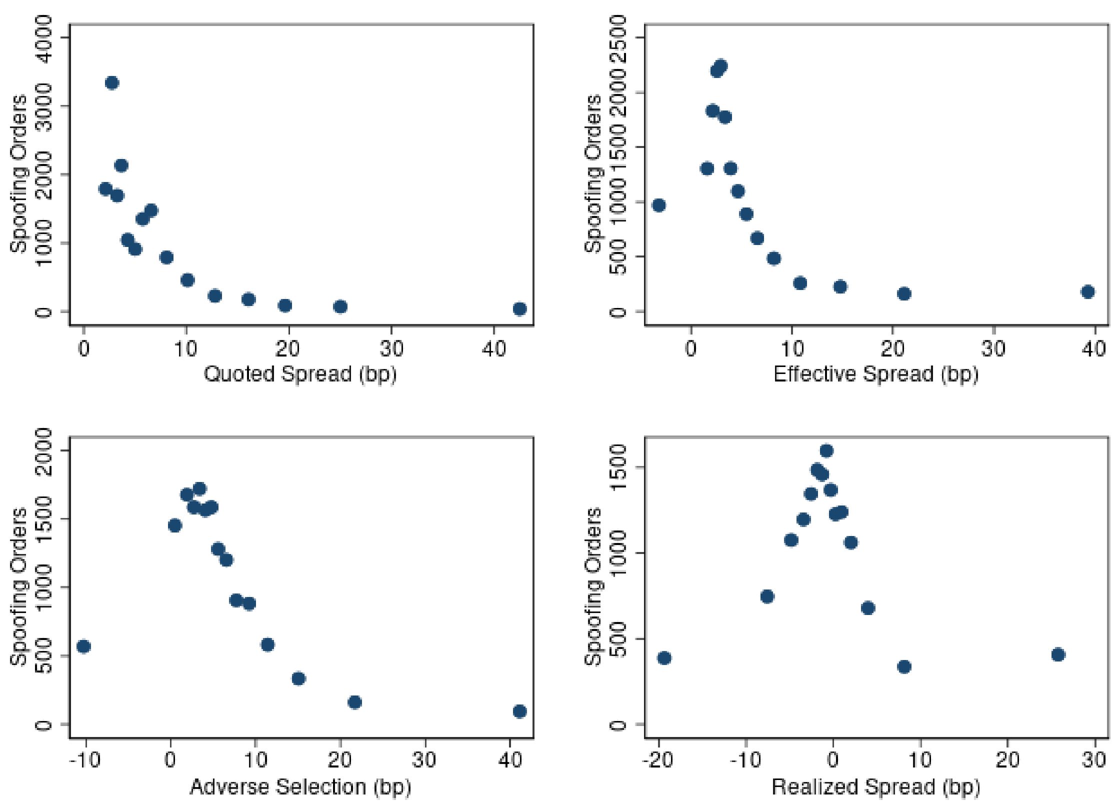


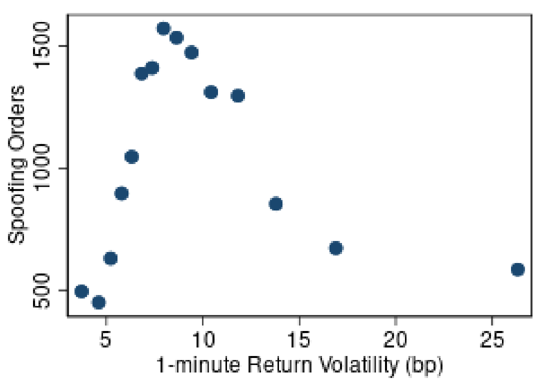
Figure 2: Spoofing and Market Quality

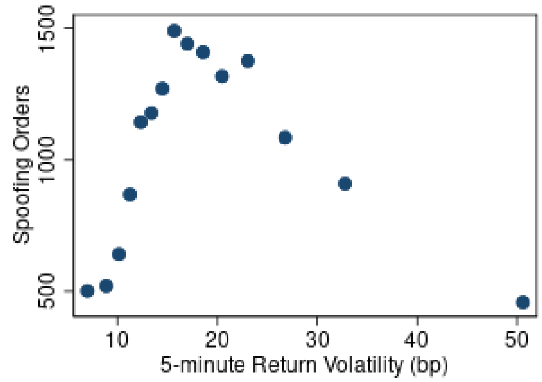
Figure 2 plots spoofing activity given lagged market quality quantiles. Spoofing is measured as the number of attempted spoofing orders for a given stock-day. Panel A plots the average level of spoofing given transaction cost quantiles, measured with time-weighted quoted spread, volume-weighted effective spread, volume-weighted adverse selection, and volume-weighted realized spread. Panel B plots the average level of spoofing given volatility quantiles, where volatility is measured with 1 and 5-minute return volatility. Panel C plots the average level of spoofing given price efficiency quantiles, where price efficiency is measured with the variance ratio and Hasbrouck (1993) pricing error σ.

Panel A: Spoofing and Lagged Transaction Costs

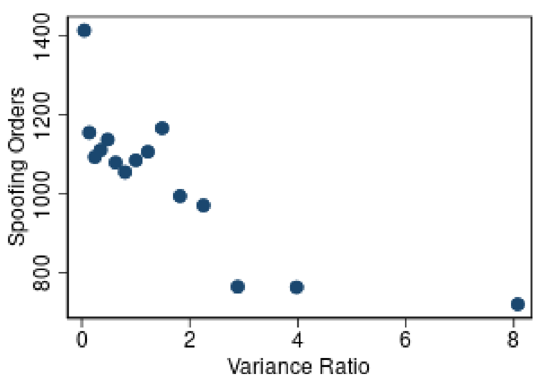


Panel B: Spoofing and Lagged Volatility





Panel C: Spoofing and Lagged Price Efficiency



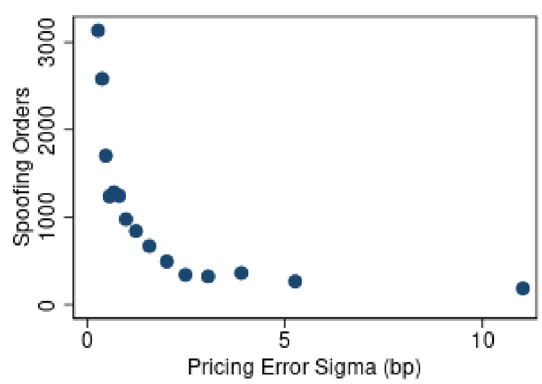


Figure 3: SEC Litigation Releases and Spoofing Activity

Figure 3 plots the average daily spoofing activity for US cross-listed and Canada-only stocks during litigation and non-litigation periods. Litigation periods are defined as the three days after a significant SEC litigation release on trade or order-based market manipulation. Spoofing is measured as the order volume from attempted spoofing scaled by daily trading volume.

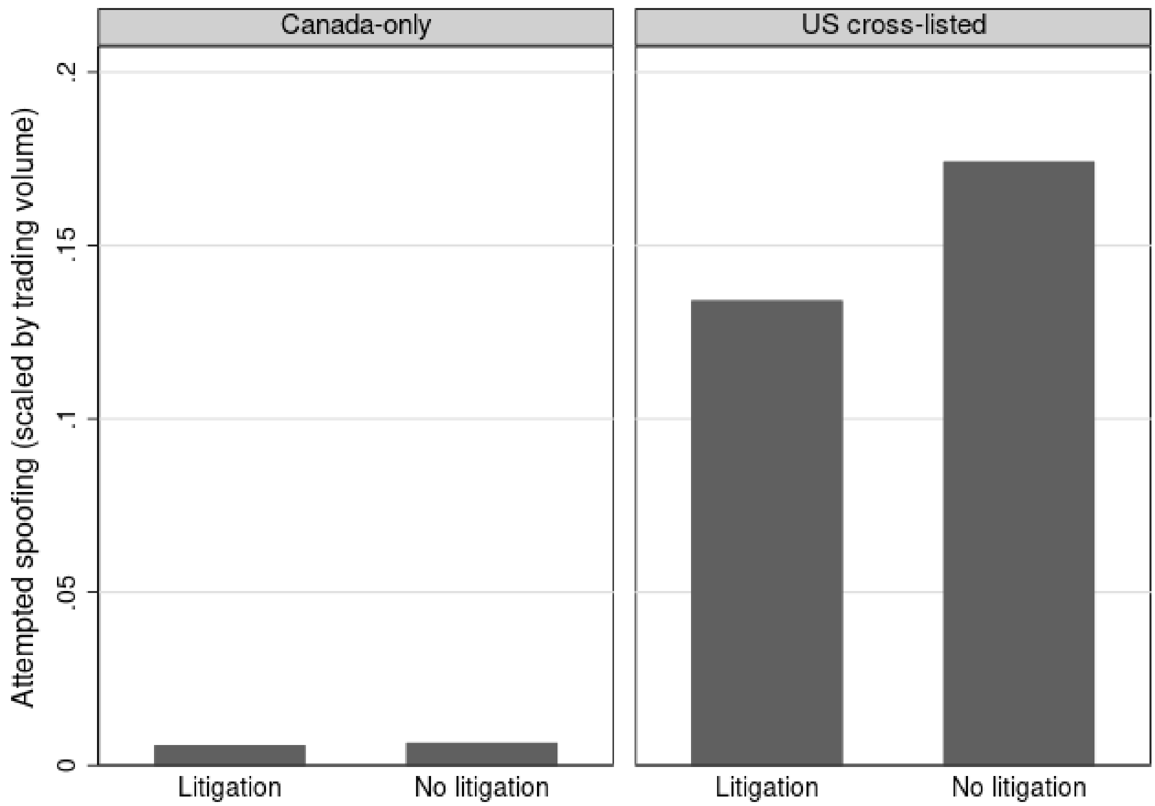


Figure 4: Dodd Frank and Spoofing Activity

Figure 4 plots the average daily spoofing activity for US cross-listed and Canada-only stocks over time. Spoofing is measured as the order volume from attempted spoofing, scaled by daily trading volume. The labels "Senate," "Committee," and "Signed into law" refer to the dates where Dodd-Frank was passed by the senate (May 20, 2010), sent to committee (June 10, 2010), and signed into law (July 21, 2010), respectively.

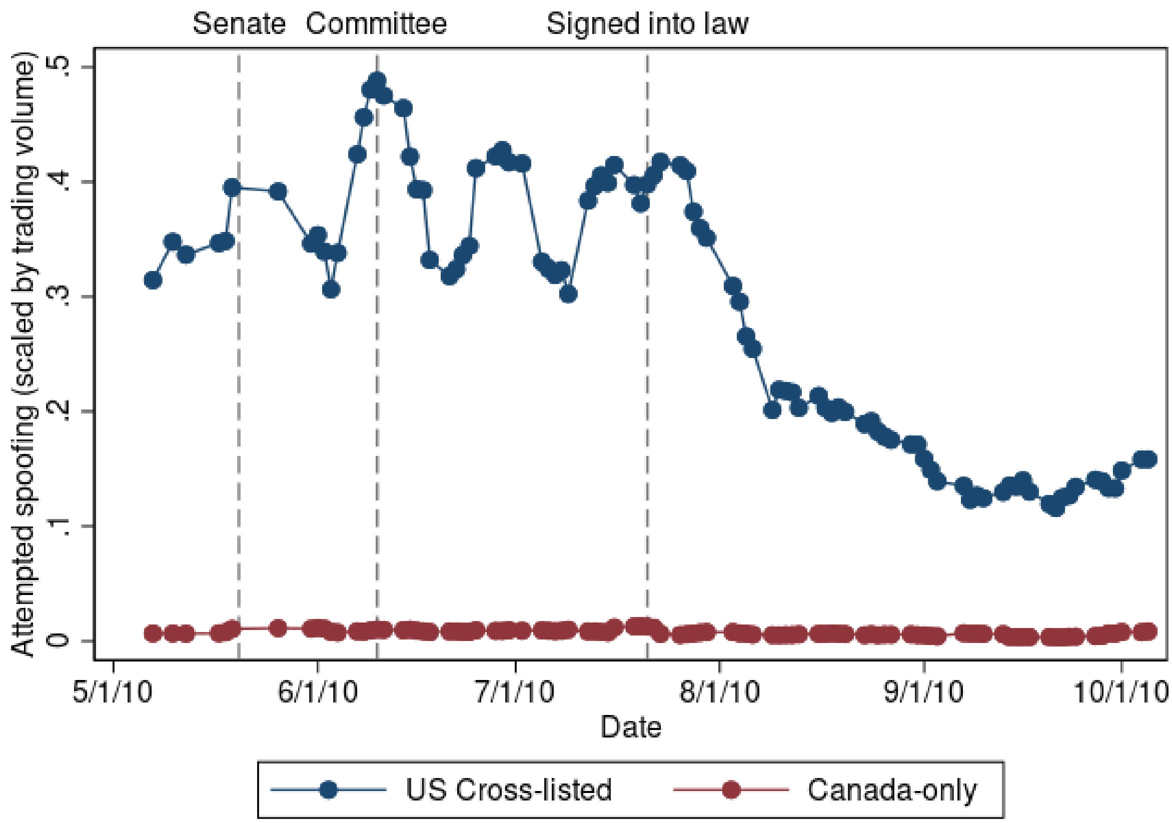


Table 1: Summary Statistics

Panel A presents stock-day level summary statistics for market quality measures. All Panel A variables except for the variance ratio, daily return, and dollar volume are reported in basis points. Panel B presents 30-minute level summary statistics for market quality measures. All Panel B variables except for variance ratio are reported in basis points. Panel C presents stock-day level summary statistics for spoofing activity. Panel D presents 30-minute level summary statistics for spoofing activity. All variables are winsorized at the 1% and 99% levels.

Panel A: Stock-Day Market Quality

	Mean	SD	p10	Median	p90	N
1-minute volatility	9.99	6.18	4.64	8.14	17.53	21415
5-minute volatility	19.49	11.98	8.97	16.04	34.14	21415
Quoted spread	12.61	13.24	2.82	7.04	28.34	21415
Effective spread	9.21	12.34	1.61	4.89	23.47	21415
Realized spread	0.25	11.35	-7.85	-0.76	9.13	21415
Adverse selection	8.91	13.11	0.43	5.73	23.47	21415
Variance ratio	1.69	2.15	0.14	1	3.93	21396
Hasbrouck σ	2.61	3.52	0.38	1.35	5.88	21306

Panel B: 30-Minute Market Quality

	Mean	SD	p10	Median	p90	N
1-minute volatility	8.86	6.63	3.12	6.92	16.91	276371
5-minute volatility	16.92	13.88	4.85	12.97	33.69	276205
Quoted spread	12.49	13.78	2.67	6.83	28.21	277561
Effective spread	9.48	14.92	1.21	4.74	25.09	276494
Realized spread	1.14	15.89	-11.76	.22	15.3	276494
Adverse selection	8.24	15.95	-2.21	4.37	24.33	276494
Variance ratio	1.24	2.16	0.07	.45	3.09	272951

Panel C: Stock-Day Spoofing

	Mean	SD	p10	Median	p90	N
Attempted spoofs (% volume)	10.1	22.06	0.00	1.59	28.22	21415
Successful spoofs (% volume)	0.18	0.25	0.00	0.08	0.52	21415
Failed spoofs (% volume)	9.92	21.91	0.00	1.43	27.78	21415
Attempted spoofs (#)	985.58	2403.68	0.00	64	2840	21415
Attempted buy spoofs (#)	486.24	1202.33	0.00	31	1376	21415
Attempted sell spoofs (#)	499.34	1252.21	0.00	31	1384	21415
Successful spoofs (#)	24.3	50.49	0.00	3	74	21415
Successful buy spoofs (#)	12.04	25.73	0.00	1	37	21415
Successful sell spoofs (#)	12.26	26.43	0.00	2	37	21415

Panel D: 30-Minute Spoofing

	Mean	SD	p10	Median	p90	N
Attempted spoofs (% volume)	0.75	1.81	0.00	0.08	1.99	277561
Successful spoofs (% volume)	0.37	.93	0.00	0.03	0.95	277561
Failed spoofs (% volume)	0.38	1.03	0.00	0.02	0.96	277561
Attempted spoofs (#)	84.1	295.91	0.00	3	195	277561
Attempted buy spoofs (#)	41.49	158.99	0.00	1	93	277561
Attempted sell spoofs (#)	42.61	169.18	0.00	1	93	277561
Successful spoofs (#)	2	6.67	0.00	0	5	277561
Successful buy spoofs (#)	0.99	3.74	0.00	0	3	277561
Successful sell spoofs (#)	1.01	3.77	0.00	0	3	277561

Table 2: Spoofing and Market Quality

Table 2 presents results of the following regression equation: $Market\ Quality_{i,t} = \beta_1 Attempted\ Spoofing_{i,t} + \boldsymbol{\beta} X + \gamma_t + \zeta_i + \epsilon_{i,t}$, where $MarketQuality_{i,t}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted adverse selection, variance ratio, or Hasbrouck (1993) pricing error σ. $Attempted\ Spoofing_{i,t}$ is the standardized attempted spoofing order volume scaled by trading volume. X represents controls for the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock and date fixed effects with γ_t and ζ_i, respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Adverse selection	(7) Variance ratio	(8) Hasbrouck σ
$Attempted\ Spoofing_{i,t}$	0.29***	0.39***	-0.16**	0.10	0.39***	-0.27***	0.03	0.05*
	(3.62)	(2.77)	(-2.44)	(1.28)	(3.28)	(-3.51)	(1.22)	(1.71)
$Average\ Dollar\ Spread_{i,t-1}$	36.94***	72.76***	92.78***	57.32***	22.17**	36.11***	4.51***	25.10***
	(7.25)	(9.06)	(6.63)	(5.08)	(2.49)	(2.82)	(2.60)	(5.57)
$Average\ Price_{i,t-1}$	-0.05***	-0.09***	-0.04**	-0.03*	-0.01	-0.02	-0.00	-0.01
	(-3.84)	(-4.07)	(-1.98)	(-1.75)	(-0.45)	(-0.94)	(-0.59)	(-1.51)
$Inverse\ Price_{i,t-1}$	42.61***	67.85***	128.69***	86.08***	74.28***	13.02	4.53***	21.16***
	(7.37)	(6.11)	(11.34)	(9.21)	(7.48)	(1.37)	(4.29)	(6.72)
$Absolute\ Return_{i,t}$	68.54***	159.86***	-7.31	39.23***	-62.87***	98.63***	-15.81***	-5.97**
	(11.87)	(13.93)	(-1.42)	(3.55)	(-4.83)	(7.40)	(-8.47)	(-2.10)
$\ln(Dollar\ Volume)_{i,t}$	1.46***	2.75***	-0.23**	-1.59***	1.10***	-2.46***	0.01	-0.19***
	(10.39)	(10.83)	(-2.31)	(-9.10)	(2.89)	(-7.23)	(0.22)	(-3.95)
$Amihud\ Illiquidity_{i,t}$	63.59	171.02*	501.90***	405.56***	-352.70***	743.50***	-19.47*	145.77***
	(1.34)	(1.88)	(5.76)	(4.04)	(-3.47)	(5.90)	(-1.69)	(4.80)
Observations	20,155	20,155	20,155	20,155	20,155	20,155	20,144	20,095
Adjusted R-squared	0.722	0.699	0.903	0.573	0.088	0.354	0.143	0.563
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 3: First Stage Litigation IV Estimate

Table 3 presents results for the following regression equation: $Attempted\ Spoofing_{i,t} = \beta_1 Litigation_t \times Treat_i + \boldsymbol{\beta} X + \gamma_t + \zeta_i + \epsilon_{i,t}$, where $Attempted\ Spoofing_{i,t}$ is the standardized attempted spoofing order volume scaled by trading volume for stock i on day t, $Litigation_t$ is an indicator variable equal to 1 if the date t is one to three days after a SEC litigation release on market manipulation, and $Treat_i$ is an indicator variable equal to 1 if stock i is cross-listed on a U.S. exchange. X represents controls for the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock and date fixed effects with γ_t and ζ_i, respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) Attempted Spoofing
$US\ Listed_i \times Litigation_t$	-0.18***
	(-5.71)
$Average\ Dollar\ Spread_{i,t-1}$	1.20
	(0.67)
$Average\ Price_{i,t-1}$	-0.01
	(-0.74)
$Inverse\ Price_{i,t-1}$	-4.96***
	(-4.81)
$Absolute\ Return_{i,t}$	-0.40
	(-0.40)
$\ln(Dollar\ Volume)_{i,t}$	-0.10**
	(-2.51)
$Amihud\ Illiquidity_{i,t}$	-3.59
	(-0.40)
Observations	20,155
Adjusted R-squared	0.526
Stock FE	Yes
Date FE	Yes

Table 4: Second Stage Litigation IV Estimate

Table 4 presents results for the following regression equation $Market\ Quality_{i,t} = \beta_1 Attempted\ \widehat{Spoofing}_{i,t} + \boldsymbol{\beta} X + \gamma_t + \zeta_i + \epsilon_{i,t}$, where $MarketQuality_{i,t}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted adverse selection, variance ratio, or Hasbrouck (1993) pricing error σ. $\widehat{Spoofing}_{i,t}$ is the predicted standardized attempted spoofing volume scaled by trading volume for stock i on day t from the first-stage IV regression in Table 3. X represents controls for the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock and date fixed effects with γ_t and ζ_i, respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Adverse selection	(7) Variance ratio	(8) Hasbrouck σ
$Attempted\ Spoofing_{i,t}$	2.50**	3.41*	0.49	3.52	-0.61	3.12	0.99*	1.36**
	(2.52)	(1.91)	(0.56)	(1.40)	(-0.27)	(1.30)	(1.67)	(1.98)
$Average\ Dollar\ Spread_{i,t-1}$	34.37***	69.23***	92.03***	53.34***	23.34**	32.17**	3.39	23.60***
	(4.51)	(6.06)	(6.47)	(3.84)	(2.37)	(2.15)	(1.19)	(4.09)
$Average\ Price_{i,t-1}$	-0.03*	-0.07***	-0.04*	-0.01	-0.01	-0.00	0.00	-0.00
	(-1.78)	(-2.66)	(-1.68)	(-0.33)	(-0.52)	(-0.02)	(0.37)	(-0.28)
$Inverse\ Price_{i,t-1}$	53.60***	82.93***	131.91***	103.11***	69.28***	29.89*	9.28***	27.68***
	(6.64)	(5.77)	(11.19)	(6.03)	(4.78)	(1.85)	(2.78)	(5.67)
$Absolute\ Return_{i,t}$	69.45***	161.11***	-7.04	40.64***	-63.28***	100.02***	-15.42***	-5.55*
	(11.80)	(14.00)	(-1.33)	(3.48)	(-4.82)	(7.13)	(-7.10)	(-1.77)
$\ln(Dollar\ Volume)_{i,t}$	1.68***	3.06***	-0.17	-1.24***	1.00**	-2.12***	0.11	-0.05
	(7.61)	(8.24)	(-1.19)	(-3.67)	(2.35)	(-4.53)	(1.26)	(-0.52)
$Amihud\ Illiquidity_{i,t}$	71.25	181.54*	504.15***	417.44***	-356.19***	755.27***	-16.07	150.65***
	(1.30)	(1.83)	(5.71)	(3.80)	(-3.40)	(5.51)	(-1.10)	(4.53)
Observations	20,155	20,155	20,155	20,155	20,155	20,155	20,144	20,095
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	32.62	32.62	32.62	32.62	32.62	32.62	32.48	32.94

Table 5: Intraday Relation Between Spoofing and Market Quality

Table 5 presents results of the following regression equation: $Market\ Quality_{i,t,j} = \beta_1 Attempted\ Spoofing_{i,t,j} + \boldsymbol{\beta}X + \theta_{it} + \phi_j + \epsilon_{i,t,j}$, where $MarketQuality_{i,t,j}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted adverse selection, or variance ratio for stock i on day t in the 30-minute interval j. $Attempted\ Spoofing_{i,t,j}$ is the standardized attempted spoofing order volume scaled by daily trading volume. X represents controls for the lag of log trading volume and absolute return. We include stock-day fixed effects with θ_{it} and 30-minute interval fixed effects with ϕ_j. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Adverse selection	(7) Variance ratio
$Attempted\ Spoofing_{i,t,j}$	1.09***	2.26***	0.10***	0.25***	-0.56***	0.84***	-0.03**
	(10.47)	(9.95)	(2.65)	(3.92)	(-3.53)	(4.53)	(-2.38)
$Volume_{i,t,j-1}$	0.18***	0.29***	0.05*	-0.09	-0.06	-0.00	0.02**
	(7.42)	(5.60)	(1.96)	(-1.27)	(-0.74)	(-0.02)	(2.34)
$Absolute\ Return_{i,t,j-1}$	178.67***	407.12***	131.26***	66.64***	-55.83**	119.25***	-11.67***
	(25.01)	(27.71)	(6.94)	(3.04)	(-2.44)	(6.22)	(-8.06)
Observations	233,598	233,449	234,603	233,719	233,719	233,719	230,562
Adjusted R-squared	0.613	0.491	0.891	0.409	0.135	0.176	0.043
Stock-day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute interval FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: The Effect of Dodd-Frank on Spoofing

Table 6 presents results for the following regression equation: $Spoofing_{i,t} = \beta_1 US\ Listed_i \times Post_t + \boldsymbol{\beta} X + \gamma_t + \zeta_i + \epsilon_{i,t}$, where $Spoofing_{i,t}$ is the count of attempted spoofing orders, count of successful spoofing orders, attempted spoofing volume scaled by trading volume, or successful spoofing volume scaled by trading volume. All spoofing variables are standardized. $Post_t$ is an indicator variable equal to 1 if the date t is on or after July 21, 2010, and $US\ Listed_i$ is an indicator variable equal to 1 if stock i is cross-listed on a U.S. exchange. X represents controls for the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock and date fixed effects with γ_t and ζ_i, respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) Count Attempted	(2) Count Successful	(3) Attempted Spoofing	(4) Successful Spoofing
$US\ Listed_i \times Post_t$	-2,076.81***	-20.07***	-1.18***	-0.66***
	(-5.85)	(-4.76)	(-6.63)	(-7.65)
$Average\ Dollar\ Spread_{i,t-1}$	-520.40	50.78	1.15	0.62
	(-0.15)	(1.11)	(0.76)	(0.61)
$Average\ Price_{i,t-1}$	-7.93	-0.39*	-0.01	-0.01*
	(-0.54)	(-1.66)	(-1.29)	(-1.90)
$Inverse\ Price_{i,t-1}$	-7,482.88***	-113.29***	-4.21***	-4.63***
	(-3.34)	(-4.50)	(-3.91)	(-6.05)
$Absolute\ Return_{i,t}$	2,793.66	82.04*	-0.12	0.77
	(1.04)	(1.80)	(-0.13)	(1.15)
$\ln(Dollar\ Volume)_{i,t}$	270.23***	11.81***	-0.12***	-0.08***
	(3.80)	(5.35)	(-3.03)	(-3.48)
$Amihud\ Illiquidity_{i,t}$	-3,396.74	259.60	-1.23	-0.54
	(-0.28)	(1.11)	(-0.15)	(-0.10)
Observations	20,155	20,155	20,155	20,155
Adjusted R-squared	0.666	0.717	0.572	0.510
Stock FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes

Table 7: Market Making Falsification

Table 7 presents results of the following regression equation: $Market\ Quality_{i,t} = \beta_1 Market\ Making_{it} + \boldsymbol{\beta} X + \gamma_t + \zeta_i + \epsilon_{i,t}$, where $MarketQuality_{i,t}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted adverse selection, variance ratio, or Hasbrouck (1993) pricing error σ. $MarketMaking_{i,t}$ is the standardized percent of orders associated with market-making activity. X represents controls for the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock and date fixed effects with γ_t and ζ_i, respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Adverse selection	(7) Variance ratio	(8) Hasbrouck σ
$Market\ Making_{i,t}$	-0.11**	-0.10	-0.14*	-0.05	-0.31**	0.27***	-0.05**	-0.06*
	(-2.49)	(-1.25)	(-1.86)	(-0.56)	(-2.53)	(2.67)	(-2.21)	(-1.80)
$Average\ Dollar\ Spread_{i,t-1}$	37.68***	73.88***	92.02***	58.81***	24.22***	36.21***	4.57***	25.50***
	(7.71)	(9.51)	(6.59)	(5.16)	(2.68)	(2.79)	(2.65)	(5.73)
$Average\ Price_{i,t-1}$	-0.05***	-0.09***	-0.04*	-0.03*	-0.01	-0.02	-0.00	-0.01
	(-3.71)	(-3.91)	(-1.91)	(-1.80)	(-0.52)	(-0.97)	(-0.56)	(-1.51)
$Inverse\ Price_{i,t-1}$	40.67***	65.23***	128.34***	86.53***	71.27***	16.51*	4.21***	20.49***
	(6.89)	(5.79)	(11.37)	(8.95)	(7.05)	(1.72)	(3.91)	(6.55)
$Absolute\ Return_{i,t}$	68.78***	160.45***	-7.13	39.21***	-62.51***	98.61***	-15.87***	-5.75**
	(11.79)	(13.91)	(-1.40)	(3.53)	(-4.83)	(7.43)	(-8.54)	(-2.04)
$\ln(Dollar\ Volume)_{i,t}$	1.41***	2.68***	-0.24**	-1.62***	1.01***	-2.41***	0.00	-0.20***
	(9.98)	(10.58)	(-2.38)	(-9.23)	(2.65)	(-7.01)	(0.07)	(-4.35)
$Amihud\ Illiquidity_{i,t}$	54.58	155.15*	501.53***	399.65***	-369.08***	748.25***	-18.78	140.96***
	(1.17)	(1.73)	(5.65)	(3.93)	(-3.67)	(5.92)	(-1.61)	(4.64)
Observations	20,140	20,140	20,140	20,140	20,140	20,140	20,132	20,086
Adjusted R-squared	0.721	0.699	0.904	0.573	0.086	0.356	0.143	0.562
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 8: Alternate Spoofing Measures Litigation First Stage

Table 8 presents results for the following regression equation: $Spoofing_{i,t} = \beta_1 SEC_t \times Treat_i + \boldsymbol{\beta} X + \gamma_t + \zeta_i + \epsilon_{i,t}$, where $Spoofing_{i,t}$ is the standardized spoofing measure for stock i on day t. Columns 1 and 2 measure spoofing with the order volume from successful spoofing and order volume from failed spoofing, respectively. Before standardization, spoofing measures are scaled by trading volume. SEC_t is an indicator variable equal to 1 if the date t is one to three days after a SEC litigation release on market manipulation, and $Treat_i$ is an indicator variable equal to 1 if stock i is cross-listed on a U.S. exchange. X represents controls for the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock and date fixed effects with γ_t and ζ_i, respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) Successful Spoofing	(2) Failed Spoofing
$US\ Listed_i \times Litigation_t$	-0.08**	-0.16***
	(-2.51)	(-5.00)
$Average\ Dollar\ Spread_{i,t-1}$	0.64	2.56
	(0.56)	(1.05)
$Average\ Price_{i,t-1}$	-0.01	-0.01
	(-1.29)	(-0.76)
$Inverse\ Price_{i,t-1}$	-5.06***	-4.59***
	(-6.98)	(-4.43)
$Absolute\ Return_{i,t}$	0.61	-0.28
	(0.87)	(-0.30)
$\ln(Dollar\ Volume)_{i,t}$	-0.07***	-0.09**
	(-2.92)	(-2.55)
$Amihud\ Illiquidity_{i,t}$	-1.85	-2.94
	(-0.30)	(-0.32)
Observations	20,155	20,155
Adjusted R-squared	0.496	0.425
Stock FE	Yes	Yes
Date FE	Yes	Yes

Table 9: Alternate Spoofing Measures Litigation Second Stage

Table 9 presents results for the following regression equation: $Market\ Quality_{i,t} = \beta_1 \widehat{Spoofing}_{i,t} + \boldsymbol{\beta}X + \gamma_t + \zeta_i + \epsilon_{i,t}$, where $MarketQuality_{i,t}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted adverse selection, variance ratio, or Hasbrouck (1993) pricing error σ. $\widehat{Spoofing}_{i,t}$ is the predicted standardized spoofing measure for stock i on day t from the first-stage IV regressions in Table 8. Panels A and B measure $\widehat{Spoofing}_{i,t}$ with instrumented standardized successful spoofing order volume and failed spoofing order volume. Both measures of spoofing are scaled by trading volume before standardization in the first stage. X represents controls for the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock and date fixed effects with γ_t and ζ_i, respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

Panel A: Successful Spoofing

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Adverse selection	(7) Variance ratio	(8) Hasbrouck σ
$Successful\ Spoofing_{i,t}$	5.84*	7.98	1.15	8.23	-1.43	7.29	2.31	3.16
	(1.91)	(1.56)	(0.54)	(1.28)	(-0.27)	(1.22)	(1.48)	(1.64)
$Average\ Dollar\ Spread_{i,t-1}$	33.64***	68.23***	91.88***	52.31***	23.52**	31.26**	3.10	23.27***
	(3.84)	(5.27)	(6.38)	(3.33)	(2.36)	(2.04)	(0.90)	(3.56)
$Average\ Price_{i,t-1}$	-0.01	-0.05	-0.03	0.02	-0.02	0.02	0.01	0.01
	(-0.46)	(-1.05)	(-1.25)	(0.31)	(-0.44)	(0.49)	(0.77)	(0.40)
$Inverse\ Price_{i,t-1}$	70.75***	106.39***	135.28***	127.30***	65.07**	51.32	16.08*	36.92***
	(4.31)	(3.84)	(9.30)	(3.61)	(2.30)	(1.55)	(1.94)	(3.54)
$Absolute\ Return_{i,t}$	64.87***	154.84***	-7.94	34.18**	-62.16***	94.30***	-17.23***	-8.18**
	(9.73)	(12.96)	(-1.51)	(2.56)	(-4.56)	(6.25)	(-6.10)	(-2.11)
$\ln(Dollar\ Volume)_{i,t}$	1.82***	3.25***	-0.14	-1.04**	0.96*	-1.94***	0.17	0.03
	(5.64)	(6.33)	(-0.74)	(-2.04)	(1.95)	(-3.36)	(1.21)	(0.15)
$Amihud\ Illiquidity_{i,t}$	73.10	184.07*	504.51***	420.04***	-356.64***	757.57***	-15.35	152.30***
	(1.17)	(1.73)	(5.72)	(3.60)	(-3.39)	(5.36)	(-0.82)	(4.23)
Observations	20,155	20,155	20,155	20,155	20,155	20,155	20,144	20,095
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	6.325	6.325	6.325	6.325	6.325	6.325	6.299	6.527

Panel B: Failed Spoofing

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Adverse selection	(7) Variance ratio	(8) Hasbrouck σ
$Failed\ Spoofing_{i,t}$	2.69**	3.68*	0.53	3.80	-0.66	3.37	1.06	1.47*
	(2.43)	(1.88)	(0.55)	(1.39)	(-0.27)	(1.29)	(1.65)	(1.95)
$Average\ Dollar\ Spread_{i,t-1}$	30.47***	63.89***	91.26***	47.83***	24.30**	27.29	1.85	21.47***
	(3.22)	(4.50)	(6.31)	(2.76)	(2.08)	(1.58)	(0.49)	(3.23)
$Average\ Price_{i,t-1}$	-0.03	-0.07**	-0.04	-0.01	-0.01	0.00	0.00	-0.00
	(-1.50)	(-2.33)	(-1.64)	(-0.22)	(-0.50)	(0.06)	(0.41)	(-0.16)
$Inverse\ Price_{i,t-1}$	53.59***	82.91***	131.91***	103.09***	69.28***	29.87*	9.28***	27.67***
	(6.62)	(5.77)	(11.19)	(6.04)	(4.78)	(1.85)	(2.80)	(5.67)
$Absolute\ Return_{i,t}$	69.20***	160.77***	-7.09	40.29***	-63.22***	99.71***	-15.52***	-5.68*
	(11.60)	(13.79)	(-1.35)	(3.46)	(-4.82)	(7.14)	(-7.24)	(-1.80)
$\ln(Dollar\ Volume)_{i,t}$	1.67***	3.05***	-0.17	-1.25***	1.00**	-2.13***	0.11	-0.06
	(7.72)	(8.36)	(-1.21)	(-3.78)	(2.37)	(-4.60)	(1.27)	(-0.56)
$Amihud\ Illiquidity_{i,t}$	70.21	180.12*	503.94***	415.98***	-355.93***	753.97***	-16.48	150.07***
	(1.24)	(1.77)	(5.70)	(3.72)	(-3.40)	(5.42)	(-1.09)	(4.41)
Observations	20,155	20,155	20,155	20,155	20,155	20,155	20,144	20,095
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	25.01	25.01	25.01	25.01	25.01	25.01	24.93	25.27

Table 10: Alternate Intraday Spoofing Measures

Table 10 presents results of the following regression equation: $Market\ Quality_{i,t,j} = \beta_1 Spoofing_{itj} + \boldsymbol{\beta} X + \theta_{it} + \phi_j + \epsilon_{i,t,j}$, where $MarketQuality_{i,j,t}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted adverse selection, or variance ratio for stock i on day t in the 30-minute interval j. Panels A and B measure spoofing with successful and failed spoofs, respectively. $Successful\ Spoofing_{i,t,j}$ is the standardized successful spoofing order volume scaled by daily trading volume. $Failed\ Spoofing_{i,t,j}$ is the standardized failed spoofing order volume scaled by daily trading volume. X represents controls for the lag of log trading volume and absolute return. We include stock-day fixed effects with θ_{it} and 30-minute interval fixed effects with ϕ_j. T-statistics are reported in parentheses and standard errors are clustered by stock.

Panel A: Successful Spoofing

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Adverse selection	(7) Variance ratio
$Successful\ Spoofing_{i,t,j}$	0.89***	1.84***	0.07**	0.29***	-0.46***	0.77***	-0.02**
	(12.27)	(12.04)	(2.10)	(4.53)	(-3.17)	(4.92)	(-2.08)
$Volume_{i,t,j-1}$	0.18***	0.30***	0.05**	-0.09	-0.06	0.00	0.02**
	(7.55)	(5.78)	(2.00)	(-1.28)	(-0.78)	(0.02)	(2.30)
$Absolute\ Return_{i,t,j-1}$	179.95***	409.90***	131.45***	66.30***	-56.52**	119.65***	-11.74***
	(25.38)	(27.97)	(6.96)	(3.03)	(-2.47)	(6.26)	(-8.06)
Observations	233,598	233,449	234,603	233,719	233,719	233,719	230,562
Adjusted R-squared	0.612	0.490	0.891	0.409	0.135	0.176	0.043
Stock-day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute interval FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Failed Spoofing

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Adverse selection	(7) Variance ratio
$Failed\ Spoofing_{i,t,j}$	0.69***	1.46***	0.07***	0.08	-0.36***	0.47***	-0.02*
	(7.53)	(6.88)	(2.76)	(1.58)	(-3.23)	(3.72)	(-1.90)
$Volume_{i,t,j-1}$	0.19***	0.31***	0.05**	-0.09	-0.07	0.01	0.02**
	(7.66)	(5.91)	(2.00)	(-1.21)	(-0.81)	(0.13)	(2.31)
$Absolute\ Return_{i,t,j-1}$	182.25***	414.51***	131.53***	67.90***	-57.68**	122.45***	-11.75***
	(25.93)	(28.43)	(6.96)	(3.10)	(-2.51)	(6.41)	(-8.09)
Observations	233,598	233,449	234,603	233,719	233,719	233,719	230,562
Adjusted R-squared	0.610	0.488	0.891	0.409	0.134	0.175	0.043
Stock-day FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute interval FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes