Does High Frequency Market Manipulation Harm Market Quality?*

Jonathan Brogaard

David Eccles School of Business,
University of Utah

Dan Li

School of Management and Economics,
Chinese University of Hong Kong, Shenzhen

Jeffrey Yang

David Eccles School of Business,
University of Utah

June 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

---

*We thank Mike Cooper, Davidson Heath, and participants at City University of Hong Kong, Chinese University of Hong Kong, and Utah for comments. Send correspondence to Jonathan Brogaard, brogaardj@eccles.utah.edu.

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. Williams and Skrzypacz (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 successful 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 average stock-day observation has 128 attempted spoofing orders, with 5 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 Williams and Skrzypacz (2021). Spoofing leads to higher return volatility, wider effective and realized spreads, 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 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 spoofing orders must be cancelled within

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

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 will undercount the true spoofing activity.

We begin the empirical analysis by documenting the prevalence and determinants of spoofing activity. In regressions of successful spoofing orders on cross-sectional and lagged market quality characteristics, we find that spoofing prevalence increases with effective spread, trading volume, and size. However, spoofing activity tends to be most prevalent in stock-days with intermediate price efficiency.

Motivated by the theoretical predictions from Williams and Skrzypacz (2021), we next focus on the relationship between spoofing and market quality. We estimate OLS panel regressions of market quality measures on successful spoofing events, while controlling for daily returns, Amihud (2002) illiquidity, the log of dollar trading volume, and stock and date fixed effects. Spoofing is statistically and economically 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 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, Williams and Skrzypacz (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 five 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 increases return volatility, raises effective and realized bid-ask spreads, increases variance ratios, and increases the Hasbrouck (1993) pricing error volatility.

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 number of attempted spoofing orders and the proportion of spoofing orders. Across the varying robustness checks the results remain economically consistent.

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. Williams and Skrzypacz (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]

[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).

We also obtain accounting data from Compustat to construct EPS excluding extraordinary items (bkvlps), book-to-market, and the natural log of market cap using closing prices from the end of calendar year 2010. For each stock, we compute these cross-sectional firm characteristics in Canadian dollars and use fiscal year 2010 data.

2.1 Liquidity 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, and Amihud (2002) illiquidity. We measure market quality with 1- and 5-minute return volatility, variance ratios, and Hasbrouck (1993) pricing error variance.

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 $2\times\frac{|\,Price-NBBO\ midpoint|}{NBBO\ midpoint}$ by the volume at each effective spread. To approximate liquidity provision revenue, we compute volume-weighted realized spreads by weighing $2\times\frac{|\,Price_t-NBBO\ midpoint_{t+5}|}{NBBO\ midpoint_t}$ by the volume at each realized spread, where $NBBO\ midpoint_{t+5}$ is the NBBO midpoint five minutes after time t. 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 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 (-1) if the trading price is greater (less) than the bid-ask average (and 0 if the trade price equals the bid-ask average), 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.

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-

[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.

spoofing order. Second, if a spoofing order is to induce a market response, it must be somewhat aggressive. We require that buy spoofing orders are greater than 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 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 several spoofing measures. First, we consider spoofing orders that are successful or unsuccessful. Successful spoofing orders are spoofing orders with executed genuine orders, while attempted spoofing orders have genuine orders that are either executed or cancelled. We also consider spoofing order volume. Percent spoofing volume is the volume of successful spoofing orders divided by total order volume. Percent attempted spoofing volume is the volume of attempted spoofing volume divided by total order volume. Lastly, we use the daily trading volume

[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.

of genuine trades divided by the total trading volume. We measure the percent variables in basis points.

Table 1 Panel B presents the stock-day level summary statistics for spoofing activity. In our sample, the average stock-day has around 5 successful spoofing orders and 128 attempted spoofing orders. However, spoofing activity is right skewed, which suggests that spoofing activity 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 valid 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.

3. Spoofing Activity

We begin by examining the determinants of spoofing activity. We regress spoofing activity on 1-day lagged market quality and liquidity measures, as well as cross sectional firm characteristics. Specifically, we estimate regressions of the following form:

$$spoofing_{i,t} = \beta_0 + \beta_1 Amihud_{i,t-1} + \beta_2 Volatility_{i,t-1} + \beta_3 VarianceRatio_{i,t-1} + \beta_4 EffectiveSpread_{i,t-1} + \beta_5 Volume_{i,t-1} + \beta_6 Return_{i,t-1} + \beta_7 HasbrouckSigma_{i,t-1} + \beta_8 EPS_i + \beta_9 BTM_i + \beta_{10} LnMKTCAP_i + \epsilon_{i,t}$$

We measure spoofing with the order volume from successful spoofing orders divided by daily order volume (reported in basis points). From our dataset of spoofing orders, we also compute the unique number of trader IDs for each stock-day that are associated with spoofing orders.

INSERT TABLE 2 ABOUT HERE

Table 2 Panel A describes the determinants of stock-day level spoofing activity. Columns (1) and (2) present results for the number of successful spoofing orders, while Columns (3) and (4) present results for the number of spoofing traders.

In all four specifications, spoofing activity has no consistent relation with lagged Amihud (2002) illiquidity, lagged 1-minute volatility, variance ratio, and lagged daily return. Lagged effective spreads positively predict spoofing and the number of spoofing traders, suggesting that spoofing is more likely to occur in less liquid stocks. In all four columns, log dollar trading volume has a positive coefficient, which suggests that spoofing activity is increasing in trading activity. Spoofing volume increases with Hasbrouck σ, which suggests that spoofing activity occurs more in stocks with worse ex-ante price efficiency. Lastly, we find that cross-sectional firm characteristics such as book-to-market and EPS do not have a consistent effect on spoofing activity and traders, while size is positively associated with both spoofing orders and spoofing traders. Because firm size may also capture unobserved liquidity characteristics, our results ultimately suggest that spoofing activity is not determined by fundamentals but rather is influenced by market quality characteristics.

Skrzypacz and Williams (2021) predict that spoofing activity should be most active in markets with moderate liquidity. As the regression specification in Table 2 does not account for a nonlinear relation between spoofing and liquidity, we compute the average percent of attempted spoofing orders (defined as attempted spoofing order volume divided by total order volume) relative to 40 liquidity or price efficiency quantiles. We measure liquidity with trading volume and effective spread and measure price efficiency with Hasbrouck σ. The results are shown in Figure 2. Panel A shows that the theoretical prediction holds at the stock-day level when proxying for liquidity with trading volume: spoofing activity is single peaked in liquidity. However, Panel B shows that spoofing generally increases in the effective spread. Panel C suggests that spoofing is most prevalent in stock-days with intermediate levels of price efficiency.

INSERT FIGURE 2 ABOUT HERE

4. Relation between Spoofing and Market Quality

Guided by the theoretical predictions in Williams and Skrzypacz (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. Table 3 presents the results for panel regressions of liquidity or price efficiency measures on spoofing and controls. For each market quality measure, we estimate regressions of the following form:

$$\ln(market\_quality_{it}) = \beta_1 Spoofing_{it} + \beta_2 Return_{it} + \beta_3 Volume_{it} + \beta_4 Amihud_{it} + \gamma_t + \zeta_i + \epsilon_{it}$$

Where $Spoofing_{it}$ is the standardized number of successful spoofing orders, $Return_{it}$ is the daily stock return, $Volume_{it}$ is the log dollar volume, and $Amihud_{it}$ is the Amihud (2002) illiquidity measure. We denote date and stock fixed effects with γ_t and ζ_i, respectively. We include several liquidity controls because the decision to spoof likely depends on a stock's liquidity (as shown in Table 2). 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 returns. 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 3 ABOUT HERE

The results in Table 3 panel A show a clear positive relation between spoofing activity and most of the inverse market quality measures. As the specification is log-linear with a standardized independent variable, the interpretation of β_1 is that a one standard deviation increase in successful spoofing orders is associated with a $100 \times \beta_1$ percent increase in the dependent variable.

Spoofing increases return volatility. We find that a one standard-deviation increase in successful spoofing orders is associated with a 3.45% and 2.67% 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. A one standard-deviation increase in successful spoofing orders is associated with a 27.81% increase in the volume-weighted effective spread and 13.05% increase in the volume-weighted realized spread. However, we find that spoofing is strongly negatively associated with quoted spreads: a one-standard deviation increase in successful spoofing orders is associated with a 2.70% decrease in the quoted spread.

Lastly, spoofing slows price discovery. A one standard-deviation increase in successful spoofing orders increases the variance ratio measure by 3.27% and Hasbrouck (1993) pricing error σ by 8.65%. 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.

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. We believe 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 6.2.

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 3 may suffer from omitted variable bias or simultaneity bias, 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, preliminary injunctions, allegations, and final judgements. The final list consists of 12 SEC releases.

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

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

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 number of spoofing orders on the interaction between $TREAT_i$, which is an indicator equal to 1 if the stock is cross-listed in the US, and SEC_t, which is an indicator equal to 1 if day t is one to five days after a SEC litigation release on market manipulation. We include controls for daily return, log of dollar trading volume, and Amihud (2002) illiquidity. We also include stock and date fixed

[9] 15 U.S. Code § 78dd

effects and cluster standard errors by stock. The first-stage results are presented in Table 4. The instrument is valid if it satisfies both the relevance and exclusion restrictions.

INSERT TABLE 4 ABOUT HERE

The first-stage results in Table 4 show that the instrument is powerful. The coefficient on $SEC_t \times Treat_i$ shows that in the five days after a SEC litigation release, US cross-listed stocks experience a 0.19 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 -5.33 and the Kleibergen-Paap rk Wald F statistic (shown in Table 5) is greater than 28. The highly significant coefficient on the instrument and large Kleibergen-Paap rk Wald F statistic suggest that the relevance condition is satisfied.

The exclusion restriction requires that $SEC_t \times Treat_i$ only affects market quality through spoofing. Threats to exclusion would have to be correlated with both $SEC_t \times Treat_i$ 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.

The second stage estimates are shown in Table 5. We regress the market quality measures from Table 3 on the predicted standardized spoofing values from the first stage estimate in Table 4. We control for daily return, log dollar volume, and Amihud (2002) illiquidity. We also include stock and date fixed effects and cluster standard errors by stock.

INSERT TABLE 5 ABOUT HERE

The results show that instrumented spoofing has a positive relation with return volatility. A one standard-deviation increase in spoofing causes a 12.33% and 8.16% increase in 1-minute and 5-minute return volatility, respectively. However, the coefficient on predicted spoofing is only statistically significant for 1-minute return volatility measure. Spoofing also has no clear relation with quoted spreads, which is reasonable given that aggressive NBBO improving spoofing orders may lower quoted spreads, while market makers who respond to spoofing may raise quoted spreads. Spoofing has a positive and economically significant relation with the variance ratio, although the relation is statistically weak. We again find strong evidence that spoofing increases effective spreads, realized spreads, and Hasbrouck (1993) pricing error σ. A one standard-deviation increase in predicted spoofing is associated with a 65%, 33%, and 26% increase in effective spread, realized spread, and Hasbrouck (1993) pricing error σ, respectively. The results suggest that spoofing harms market quality.

6. Robustness

We apply a battery of robustness tests to validate our spoofing measures and ensure that our results are not driven by market-making activity or our choice of spoofing measure.

6.1 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, we believe that 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, it may not be appropriate to use the law's passage date, July 21, 2010, to indicate the "post" period in the difference-in-differences analysis. 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 3. Panels A, B, and C use the number of successful spoofing orders, order volume associated with successful spoofing orders, and order volume associated with attempted spoofing orders, respectively.

INSERT FIGURE 3 ABOUT HERE

The three definitions of spoofing yield a similar time-series pattern. From the start of the sample to May 20, spoofing in US cross-listed and Canada-only stocks is flat across time in Panel A and C and is declining in Panel B. 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. However, after Dodd-Frank entered the conference committee stage on June 10, spoofing falls again and remains at the lower level. The passage of Dodd-Frank on July 21 has no noticeable change in spoofing activity. Because spoofing starts to have a noticeable change (in US cross-listed relative to Canada-only stocks) after the senate passed Dodd-Frank, we use May 20, 2010 to define the start of the post period in the difference-in-differences analysis. However, our results are robust to using the conference committee date to indicate the post period.

We verify formally that spoofing declines in Table 6. We regress the three spoofing measures on a $Treat_i \times Post_t$ variable that is equal to 1 if stock i is cross-listed in the US and if day t is on or after May 20, 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 successful spoofing orders falls by 1.26 standard deviations, while the percent of order volume associated with successful and attempted spoofing falls by 0.85 and 1.48 standard deviations, respectively. 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.

6.2 **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 3 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 consistent with the existing literature that increased algorithmic trading improves liquidity (Hendershott, Jones, and Menkveld, 2011; Brogaard, Hendershott, and Riordan, 2014). These results are also the opposite of what we find for spoofing activity, which suggests that the spoofing measures are likely not capturing genuine market-making activity.

6.3 **Alternative Spoofing Definitions**

The main results measure spoofing as the standardized number of successful spoofing orders. We test alternative definitions of spoofing in this section. Namely, we use the standardized number of attempted spoofing orders and the proportion of total order volume associated with attempted spoofing orders.

We define the number of attempted spoofing orders as all spoofing orders that pass all six filters from Section 2.2, but do not require that the associated genuine orders are executed. Thus, attempted spoofing orders contain both the successful and unsuccessful spoofing orders. We have shown in previous sections that successful spoofing strategies tend to harm market quality. However, it is also likely that attempted spoofing strategies do too.

We measure the proportion of attempted spoofing order volume to total order volume by summing attempted spoofing order volume and dividing by total order volume for each stock-day. Measuring spoofing in levels affords relatively clean measurement, but a valid concern is that spoofing levels are too small to plausibly affect market quality. By measuring spoofing in proportions, we can examine how market quality responds to spoofing even after adjusting for the total level of order volume. For each alternative spoofing measure, we re-estimate the SEC litigation release IV 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 $SEC_t \times Treat_i$ 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 statistics (shown in Table 9) are over 30 for both alternate spoofing measures suggesting that the first stage estimates are powerful.

INSERT TABLE 8 ABOUT HERE

Table 9 presents the second stage estimates for the SEC litigation IV. Panel A presents results for attempted spoofing orders, while Panel B presents results for the percent of spoofing order volume. Consistent with the SEC litigation IV results in Table 5, we find a positive relation between spoofing and volatility and no statistically significant relation between spoofing and quoted spread. We again find a weak positive relation between instrumented spoofing and the variance ratio, and strong statistically positive relations between spoofing and effective spread, realized spread, and Hasbrouck (1993) pricing error σ. The coefficients across spoofing measures are stable and suggest that the IV results are not spurious.

INSERT TABLE 9 ABOUT HERE

7. **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 Williams and Skrzypacz (2021), spoofing increases return volatility, increases effective and realized spreads, and slows price discovery.

We develop a tractable six-step filtering process to identify spoofing orders and study the prevalence of spoofing. We show that spoofing activity can be predicted by some ex-ante market-quality variables and not by firm fundamentals. Consistent with Williams and Skrzypacz (2021), spoofing activity is single-peaked in liquidity when measured with trading volume and Hasbrouck (1993) pricing error σ. However, spoofing is most prevalent in stocks with the highest effective spreads.

OLS regressions show that on average, spoofing activity is associated with worse market quality. SEC Litigation Releases, we exploit the variation in spoofing in US-Canada cross-listed and Canada-only stocks in a difference-in-differences framework to provide causal evidence that spoofing harms market quality. The reduction in spoofing decreases volatility, increases decreases costs, and improves price efficiency.

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 Williams and Skrzypacz (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.

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	
Successful spoofs	Number of successful spoofing orders as defined by the procedure in Section 2.2. A successful spoof must have an associated genuine order that is also executed.
Attempted spoofs	Number of attempted spoofing orders as defined by the procedure in Section 2.2. Includes both successful and unsuccessful spoofs, meaning that the associated genuine order does not have to be executed.
Percent spoofing volume	The order volume from successful spoofing orders divided by the total daily order volume.
Percent attempted spoofing volume	The order volume associated with attempted spoofing orders divided by the total daily order volume.
Percent genuine spoofing trades	The trading volume associated with genuine orders divided by total daily trading volume. Genuine orders are defined in Section 2.2 and are the legitimate orders placed on the opposite side of spoofing orders.
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 $2 \times \frac{\lvert Price-NBBO\ midpoint\rvert}{NBBO\ midpoint}$.
Realized spread	Volume-weighted realized spread, where each realized spread is $2 \times \frac{\lvert Price_t-NBBO\ midpoint_{t+5}\rvert}{NBBO\ midpoint_t}$.
Variance ratio	Lo and MacKinlay (1988) variance ratios using 1 and 30-minute return variances: $\lvert 1-30 \times \frac{Var_{1\ minute}(ret)}{Var_{30\ minute}(ret)}\rvert$.
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 greater (less) than the bid-ask average (and 0 if the trade price equals the bid-ask average), 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 one-way trading volume.
Order volume	Total order volume.

Daily return	Percent 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.
Fundamental characteristics	
EPS	Diluted EPS excluding extraordinary items (epsfx) using earliest fiscal year data after January 1, 2010 and before December 31, 2012.
Book-to-market	Book value per share (bkvlps) divided by price per share using earliest fiscal year data after January 1, 2010 and before December 31, 2012.
Ln(Market cap)	ln $(price \times shares\ outstanding)$ using earliest fiscal year data after January 1, 2010 and before December 31, 2012.

Figure 1: Spoofing Example

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

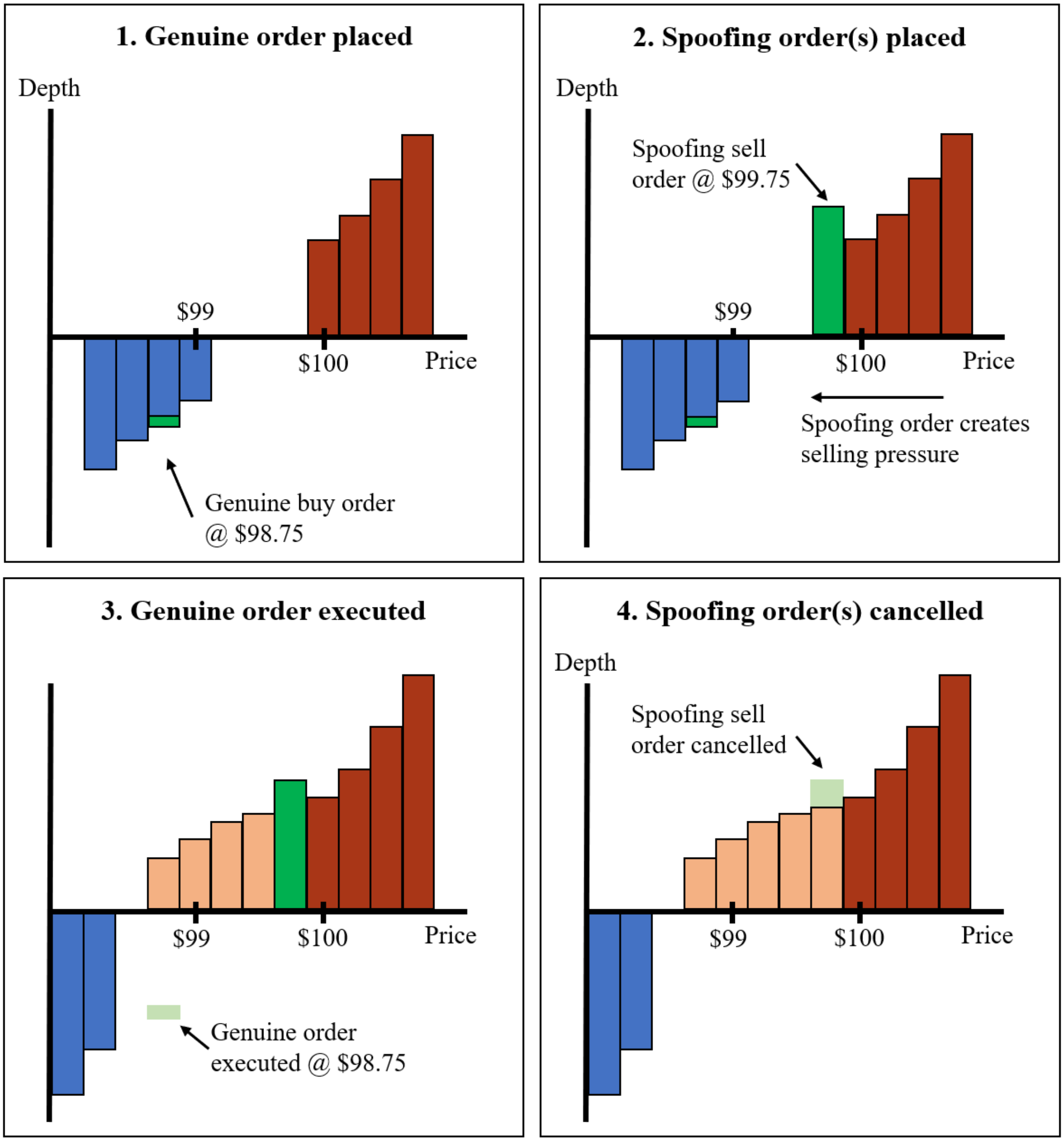
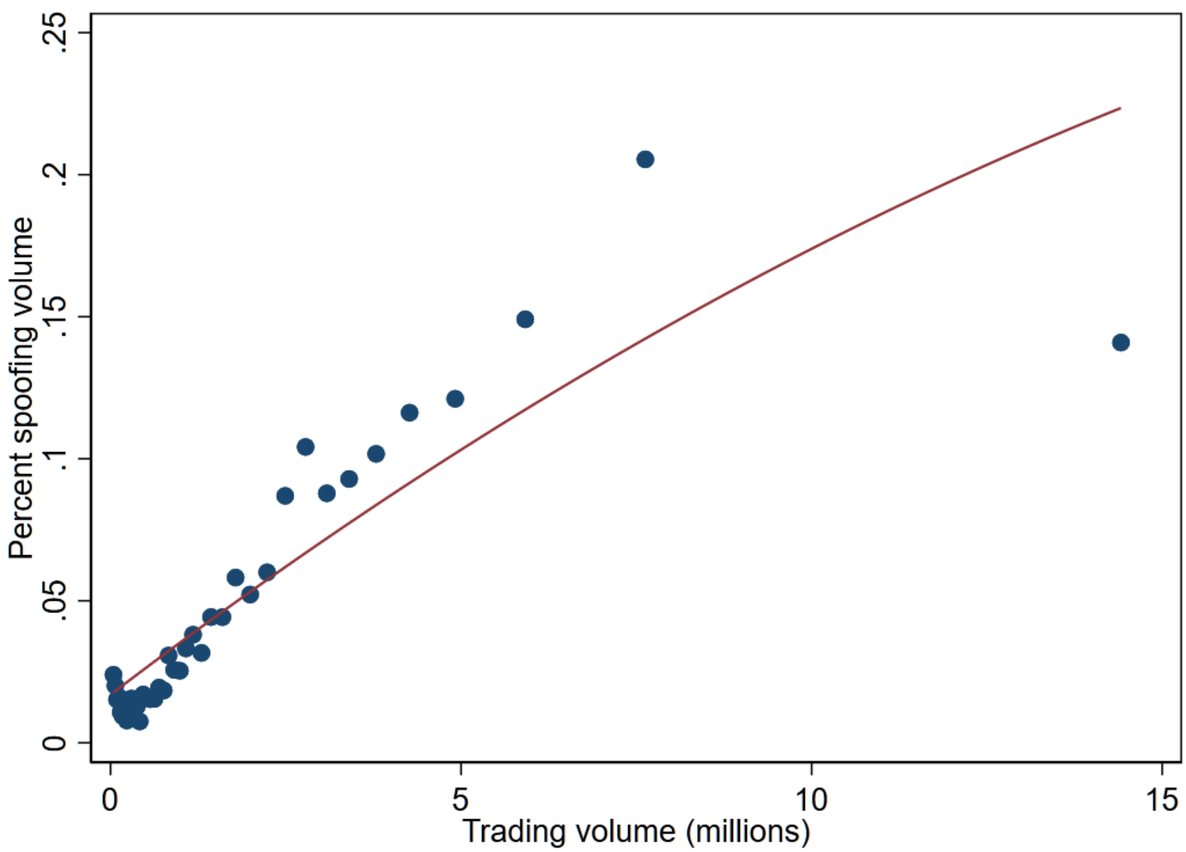


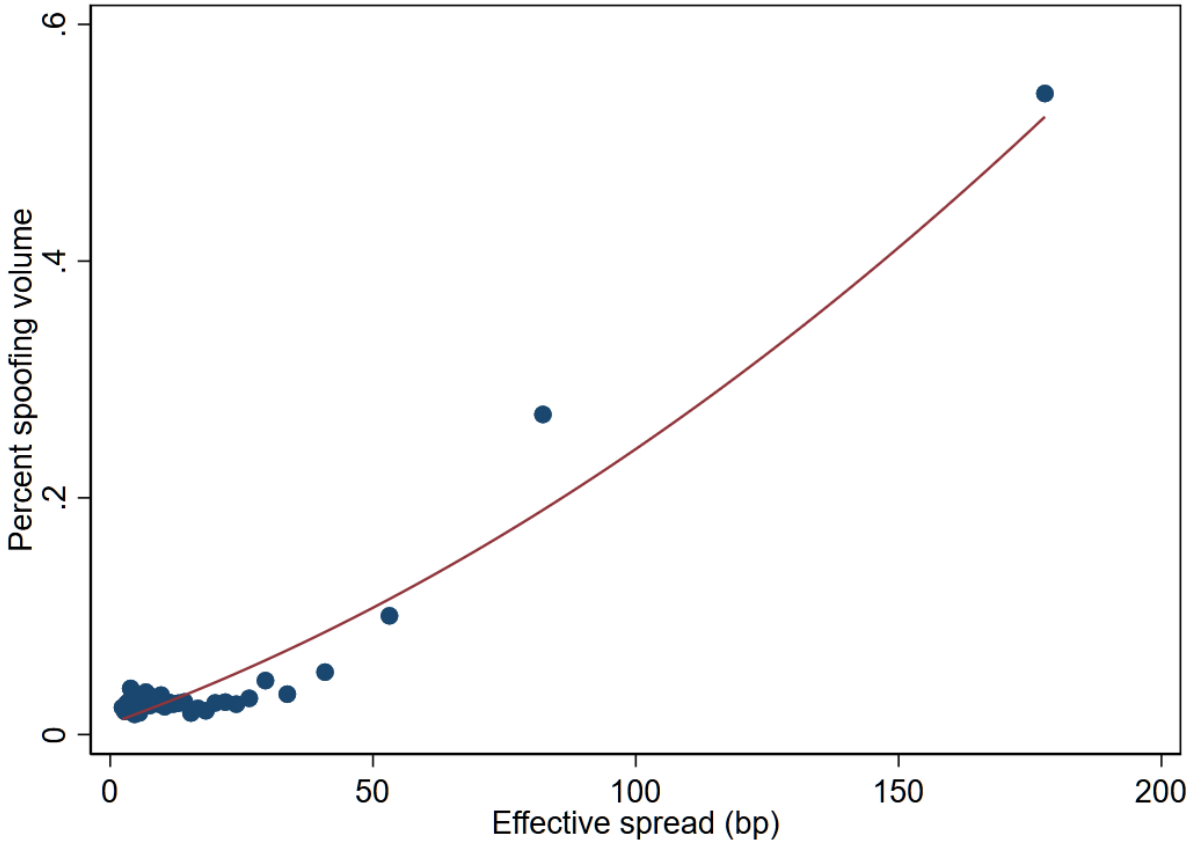
Figure 2: Spoofing and Market Quality

Figure 2 plots the average percentage of daily attempted spoofing orders for 40 liquidity or price efficiency quantiles. Panel A defines liquidity as the lagged trading volume (in million), Panel B defines liquidity as the effective spread (bp), and Panel C defines price efficiency as the Hasbrouck pricing error σ (bp).

Panel A: Trading Volume



Panel B: Effective Spread



Panel C: Hasbrouck σ

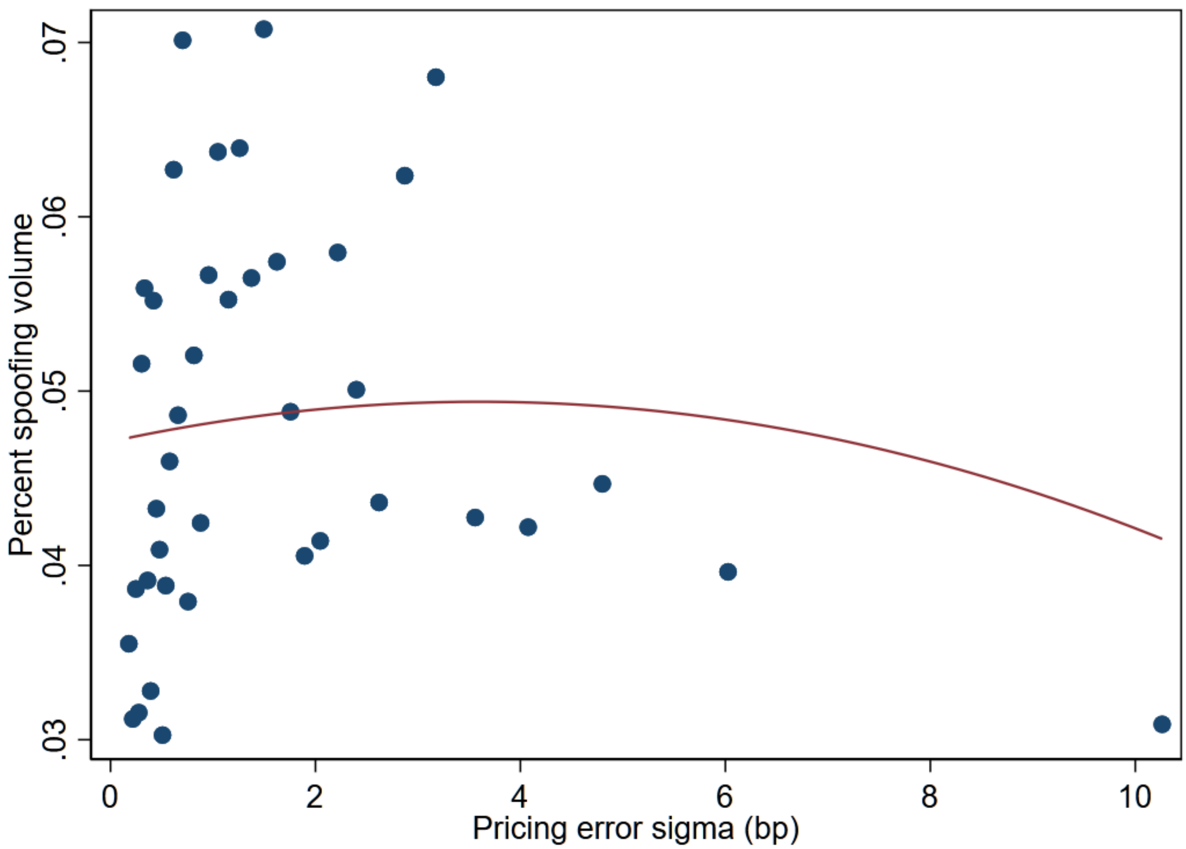
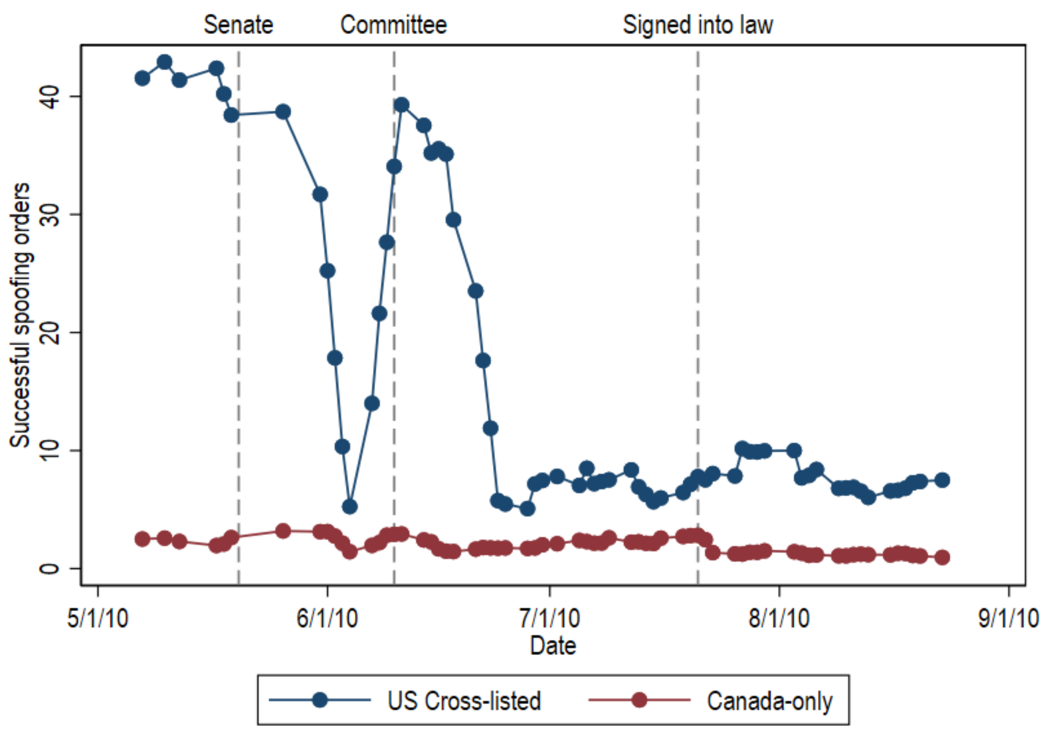


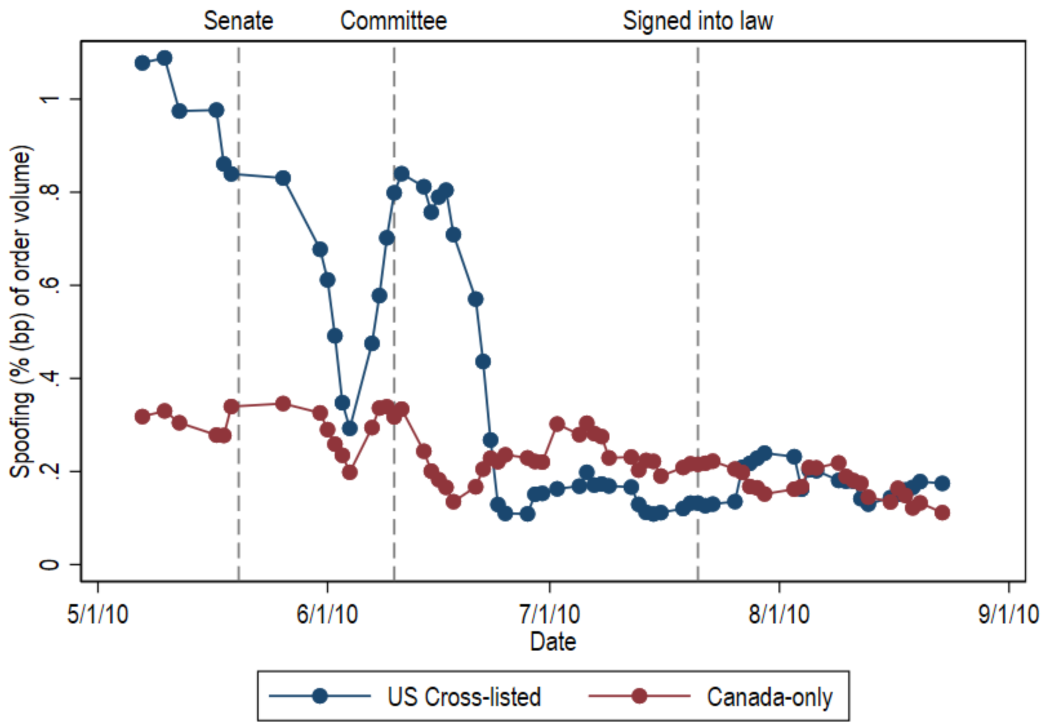
Figure 3: Dodd Frank and Spoofing Activity

Figure 3 plots the average daily spoofing activity for US cross-listed and Canada-only stocks. Panel A shows the average number of successful spoofing orders, Panel B shows the percent of order volume associated with successful spoofing, and Panel C shows the Percent of order volume associated with attempted spoofing. 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.

Panel A: Successful spoofing orders



Panel B: Percent of order volume associated with successful spoofing



Panel C: Percent of order volume associated with attempted spoofing

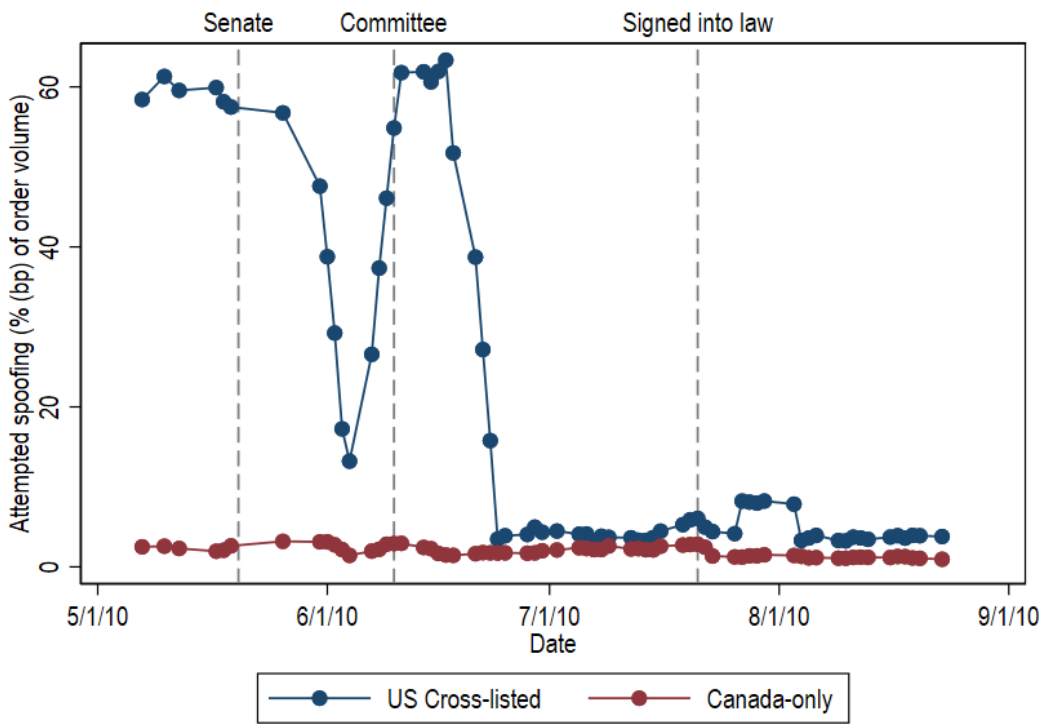


Table 1: Summary Statistics

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 stock-day level summary statistics for spoofing activity. All variables are winsorized at the 1% and 99% levels.

Panel A: Market Characteristics

	Mean	SD	p10	Median	p90	N
1-minute return volatility (bp)	9.83	5.91	4.61	8.07	17.26	20603
5-minute return volatility (bp)	19.20	11.51	8.91	15.91	33.63	20603
Quoted spread (bp)	91.91	80.14	22.90	64.93	200.24	20603
Effective spread (bp)	20.41	32.47	3.59	9.48	43.04	20603
Realized spread (bp)	40.39	35.65	14.67	29.59	74.74	20603
Variance ratio	1.68	2.10	0.14	1.00	3.91	20585
Hasbrouck pricing error σ (bp)	1.85	2.27	0.30	1.01	4.23	20504
p90 Hasbrouck error (bp)	2.72	3.15	0.47	1.51	6.31	20504
p100 Hasbrouck error (bp)	16.85	22.47	3.26	9.58	35.72	20504
Daily return (%)	0.05	1.75	-1.90	0.00	2.03	20603
Dollar trading volume (millions)	52.33	71.92	1.80	20.07	154.80	20603

Panel B: Spoofing Activity

	Mean	SD	p10	Median	p90	N
Successful spoofs	4.65	12.69	0.00	0	13	20603
Successful buy spoofs	2.20	6.52	0.00	0	6	20603
Successful sell spoofs	2.45	7.74	0.00	0	6	20603
Attempted spoofs	127.86	527.11	0.00	4	209	20603
Attempted buy spoofs	57.53	251.49	0.00	2	94	20603
Attempted sell spoofs	70.34	343.91	0.00	2	97	20603
Percent spoofing volume (bp)	.20	.52	0.00	0.00	.55	20603
Percent attempted spoofing volume (bp)	4.84	20.56	0.00	.49	5.97	20603
Percent genuine spoofing trades (bp)	10.2	29.59	0.00	0.00	25.25	20603

Table 2: Spoofing Characteristics

Table 2 describes spoofing activity in the sample in regressions of the form: $spoofing_{i,t} = \beta X_{i,t-1} + \epsilon_{i,t}$, where $X_{i,t-1}$ is a vector of 1-day lagged independent variables and $spoofing_{i,t}$ is either the daily spoofing order volume divided by total order volume or the number of traders that spoof for stock i on day t. Standard errors are clustered by stock.

	(1) Spoofing	(2) Spoofing	(3) Traders	(4) Traders
Amihud illiquidity	-5.00**	-2.67*	51.17	31.06
	(-2.12)	(-1.73)	(1.29)	(0.76)
1-minute return volatility	-0.00	-0.01	2.31***	4.76***
	(-0.08)	(-0.18)	(2.72)	(5.31)
Effective spread	0.19***	0.17***	2.74***	3.17***
	(9.06)	(7.63)	(6.04)	(8.24)
Dollar trading volume	0.12***	0.08***	3.52***	2.10***
	(5.46)	(4.75)	(9.87)	(5.08)
Daily return	-0.13	-0.01	-6.57*	-0.47
	(-0.74)	(-0.04)	(-1.82)	(-0.11)
Variance ratio	0.01**	-0.00	-0.17**	-0.13**
	(2.05)	(-0.42)	(-2.02)	(-2.26)
Hasbrouck σ	0.04	0.04**	-2.58***	-1.97***
	(1.48)	(2.21)	(-5.02)	(-4.39)
Earnings per share		-0.03**		0.22
		(-2.50)		(1.02)
Book to market		0.14***		-0.74
		(2.82)		(-0.63)
Ln(Market cap)		0.07***		2.36***
		(4.05)		(5.00)
Observations	18,939	16,672	18,939	16,672
Adjusted R-squared	0.136	0.170	0.382	0.427

Table 3: Spoofing and Market Quality

Table 3 presents results of the following regression equation: $\ln(MarketQuality_{it}) = \beta_1 Spoofing_{it} + \beta_2 Return_{it} + \beta_3 Volume_{it} + \beta_4 Amihud_{it} + \gamma_t + \zeta_i + \epsilon_{it}$, where $MarketQuality_{it}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, variance ratio, or Hasbrouck (1993) pricing error σ. $Spoofing_{it}$ is the standardized number of successful spoofing orders, $Return_{it}$ is the daily stock return, $Volume_{it}$ is log dollar trading volume, and $Amihud_{it}$ is the Amihud (2002) illiquidity measure. 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) Variance ratio	(7) Hasbrouck σ
Spoofing orders	0.0345***	0.0267***	-0.0270***	0.2781***	0.1305***	0.0327*	0.0865***
	(7.09)	(5.73)	(-3.56)	(15.66)	(15.42)	(1.87)	(7.94)
Daily return	0.1865	0.6067***	0.0173	-0.1270	0.5853***	-4.1991***	-0.1761
	(1.23)	(3.70)	(0.06)	(-0.48)	(2.84)	(-7.82)	(-0.88)
Dollar trading volume	0.1390***	0.1510***	-0.0216**	-0.0173	0.1348***	-0.1458***	-0.1569***
	(15.89)	(17.78)	(-2.07)	(-1.34)	(11.68)	(-6.49)	(-12.16)
Amihud illiquidity	14.5592***	16.8269***	5.5861***	10.6168***	16.4908***	-24.3194***	6.9889***
	(11.28)	(13.50)	(3.89)	(7.07)	(11.01)	(-10.28)	(4.84)
Observations	20,597	20,597	20,597	20,597	20,597	20,580	20,500
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Within R-squared	0.0990	0.0983	0.00338	0.157	0.103	0.0100	0.0520

Table 4: First Stage Litigation IV Estimate

Table 4 presents results for the following regression equation: $Spoofing_{it} = \beta_1 SEC_t \times Treat_i + \beta_2 Amihud_{it} + \beta_3 Ret_{it} + \beta_4 Volume_{it} + \gamma_t + \zeta_i + \epsilon_{it}$, where $Spoofing_{it}$ is the standardized number of successful spoofing orders for stock i on day t, SEC_t is an indicator variable equal to 1 if the date t is one to five 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. $Amihud_{it}$ is the Amihud (2002) illiquidity measure, Ret_{it} is daily stock return, and $Volume_{it}$ is the log dollar trading volume. 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) Spoofing orders
$SEC_t \times Treat_i$	-0.1894***
	(-5.33)
Daily return	0.5602
	(1.61)
Dollar trading volume	0.1766***
	(5.57)
Amihud illiquidity	-0.9680
	(-0.84)
Observations	20,597
Adjusted R-squared	0.422
Stock FE	Yes
Date FE	Yes

Table 5: Second Stage Litigation IV Estimate

Table 5 presents results for the following regression equation: $\ln(MarketQuality_{it}) = \beta_1 \widehat{Spoofing}_{it} + \beta_2 Return_{it} + \beta_3 Volume_{it} + \gamma_t + \zeta_i + \epsilon_{it}$, where $MarketQuality_{it}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, variance ratio, or Hasbrouck (1993) pricing error σ. $\widehat{Spoofing}_{it}$ is the predicted standardized number of successful spoofing orders for stock i on day t from the first-stage IV regression in Table 4, $Return_{it}$ is the daily stock return, $Volume_{it}$ is log dollar trading volume, and $Amihud_{it}$ is the Amihud (2002) illiquidity measure. 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) Variance ratio	(7) Hasbrouck σ
$\widehat{\text{Spoofing}}$	0.1233**	0.0816	0.0833	0.6538***	0.3335***	0.4540*	0.2568***
	(2.08)	(1.30)	(0.80)	(6.35)	(3.49)	(1.76)	(2.81)
Daily return	0.1351	0.5749***	-0.0466	-0.3446	0.4678**	-4.4424***	-0.2777
	(0.82)	(3.30)	(-0.15)	(-1.17)	(2.12)	(-7.42)	(-1.24)
Dollar trading volume	0.1231***	0.1412***	-0.0413*	-0.0845***	0.0986***	-0.2211***	-0.1879***
	(8.82)	(10.19)	(-1.72)	(-3.08)	(4.38)	(-4.13)	(-8.85)
Amihud illiquidity	14.6408***	16.8773***	5.6874***	10.9619***	16.6773***	-23.9348***	7.1308***
	(11.28)	(13.55)	(3.88)	(7.14)	(11.16)	(-9.68)	(4.86)
Observations	20,597	20,597	20,597	20,597	20,597	20,580	20,500
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	28.42	28.42	28.42	28.42	28.42	28.38	28.87

Table 6: The Effect of Dodd-Frank on Spoofing

Table 6 presents results for the following regression equation: $Spoofing_{it} = \beta_1 Post_t \times Treat_i + \beta_2 Amihud_{it} + \beta_3 Ret_{it} + \beta_4 Volume_{it} + \gamma_t + \zeta_i + \epsilon_{it}$, where $Spoofing_{it}$ is the standardized number of successful spoofing, standardized percent of order volume associated with successful spoofing orders, and the standardized percent of order volume associated with attempted (successful and failed) spoofing orders. $Post_t$ is an indicator variable equal to 1 if the date t is on or after May 20, 2010, and $Treat_i$ is an indicator variable equal to 1 if stock i is cross-listed on a U.S. exchange. $Amihud_{it}$ is the Amihud (2002) illiquidity measure, Ret_{it} is daily stock return, and $Volume_{it}$ is the log dollar trading volume. 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)	(2)	(3)
	Spoofing orders	Percent of order volume associated with successful spoofing	Percent of order volume associated with attempted spoofing
$Post_t \times Treat_i$	-1.2582***	-0.8467***	-1.4787***
	(-6.28)	(-3.87)	(-6.42)
Daily return	0.6581	-0.1499	0.4523
	(1.32)	(-0.25)	(0.83)
Dollar trading volume	0.1825***	0.2449***	0.0739**
	(4.85)	(5.68)	(2.03)
Amihud illiquidity	3.6874**	1.0342	1.9756
	(2.34)	(0.39)	(1.07)
Observations	7,919	7,919	7,919
Stock FE	Yes	Yes	Yes
Date FE	Yes	Yes	Yes
Within R-squared	0.0767	0.0399	0.0741

Table 7: Falsification

Table 7 presents results of the following regression equation: $\ln(MarketQuality_{it}) = \beta_1 MarketMaking_{it} + \beta_2 Return_{it} + \beta_3 Volume_{it} + \beta_4 Amihud_{it} + \gamma_t + \zeta_i + \epsilon_{it}$, where $MarketQuality_{it}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, variance ratio, or Hasbrouck (1993) pricing error σ. $MarketMaking_{it}$ is the standardized percent of orders associated with market-making activity, $Return_{it}$ is the daily stock return, $Volume_{it}$ is log dollar trading volume, and $Amihud_{it}$ is the Amihud (2002) illiquidity measure. 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) Variance ratio	(7) Hasbrouck σ
Market-making	-0.0117***	-0.0129***	-0.0139**	-0.0286***	-0.0054	-0.0095	-0.0412***
	(-3.24)	(-3.27)	(-2.56)	(-3.72)	(-1.00)	(-0.69)	(-4.87)
Daily return	0.1786	0.5913***	-0.0315	-0.0341	0.6481***	-4.2030***	-0.2223
	(1.18)	(3.60)	(-0.11)	(-0.12)	(3.07)	(-7.88)	(-1.11)
Dollar trading volume	0.1454***	0.1560***	-0.0261**	0.0330**	0.1583***	-0.1397***	-0.1402***
	(16.81)	(18.41)	(-2.57)	(2.16)	(12.94)	(-6.31)	(-10.36)
Amihud illiquidity	14.5099***	16.7829***	5.5900***	10.3184***	16.3629***	-24.3639***	6.8530***
	(11.21)	(13.41)	(3.86)	(6.58)	(10.72)	(-10.30)	(4.77)
Observations	20,597	20,597	20,597	20,597	20,597	20,580	20,500
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Within R-squared	0.0901	0.0944	0.00242	0.00624	0.0447	0.00968	0.0358

Table 8: Alternate Spoofing Measures Litigation First Stage

Table 8 presents results for the following regression equation: $Spoofing_{it} = \beta_1 SEC_t \times Treat_i + \beta_2 Amihud_{it} + \beta_3 Ret_{it} + \beta_4 Volume_{it} + \gamma_t + \zeta_i + \epsilon_{it}$, where $Spoofing_{it}$ is the standardized spoofing measure for stock i on day t. Columns 1 and 2 measure $Spoofing_{it}$ with standardized attempted spoofs and the percent of attempted spoofing order volume, respectively. 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. $Amihud_{it}$ is the Amihud (2002) illiquidity measure, Ret_{it} is daily stock return, and $Volume_{it}$ is the dollar trading volume. 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 orders	(2) Percent of order volume associated with spoofing
$SEC \times Treat$	-0.28***	-0.28***
	(-6.07)	(-6.23)
Daily return	0.69*	0.40
	(1.94)	(1.18)
Dollar trading volume	0.08***	0.08***
	(3.40)	(3.31)
Amihud illiquidity	-4.29***	-4.43***
	(-2.88)	(-2.73)
Observations	20,597	20,597
Stock FE	Yes	Yes
Date FE	Yes	Yes
Within R-squared	0.00831	0.00792

Table 9: Alternate Spoofing Measures Litigation Second Stage

Table 9 presents results for the following regression equation: $\ln(MarketQuality_{it}) = \beta_1 \widehat{Spoofing}_{it} + \beta_2 Return_{it} + \beta_3 Volume_{it} + \gamma_t + \zeta_i + \epsilon_{it}$, where $MarketQuality_{it}$ is 1-minute return volatility, 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, variance ratio, or Hasbrouck (1993) pricing error σ. $\widehat{Spoofing}_{it}$ is the predicted standardized spoofing measure for stock i on day t from the first-stage IV regression in Table 8. Panels A and B measure $\widehat{Spoofing}_{it}$ with instrumented standardized attempted spoofing orders and percent of attempted spoofing order volume, respectively. $Return_{it}$ is the daily stock return, $Volume_{it}$ is log dollar trading volume, and $Amihud_{it}$ is the Amihud (2002) illiquidity measure. 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 presents results for attempted spoofing orders, while Panel B presents results for the percent of order volume associated with spoofing.

Panel A: Attempted Spoofing Orders

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Variance ratio	(7) Hasbrouck σ
$\widehat{\text{Spoofing}}$	0.0831**	0.0550	0.0561	0.4406***	0.2248***	0.3058*	0.1747***
	(2.07)	(1.29)	(0.78)	(7.41)	(3.54)	(1.81)	(2.89)
Daily return	0.1472	0.5829***	-0.0384	-0.2802	0.5006**	-4.3980***	-0.2507
	(0.92)	(3.37)	(-0.13)	(-1.02)	(2.32)	(-7.50)	(-1.15)
Dollar trading volume	0.1381***	0.1511***	-0.0312**	-0.0052	0.1390***	-0.1659***	-0.1564***
	(14.97)	(16.44)	(-2.32)	(-0.37)	(10.10)	(-6.17)	(-11.59)
Amihud illiquidity	14.8776***	17.0339***	5.8472***	12.2175***	17.3177***	-23.0633***	7.6399***
	(11.34)	(13.47)	(3.88)	(7.78)	(11.20)	(-9.20)	(5.19)
Observations	20,597	20,597	20,597	20,597	20,597	20,580	20,500
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	36.89	36.89	36.89	36.89	36.89	36.88	36.70

Panel B: Percent of order volume associated with attempted spoofing

	(1) 1-minute volatility	(2) 5-minute volatility	(3) Quoted spread	(4) Effective spread	(5) Realized spread	(6) Variance ratio	(7) Hasbrouck σ
$\widehat{\text{Spoofing}}$	0.0829**	0.0548	0.0559	0.4393***	0.2241***	0.3050*	0.1739***
	(2.06)	(1.29)	(0.78)	(7.53)	(3.57)	(1.84)	(2.89)
Daily return	0.1711	0.5987***	-0.0223	-0.1538	0.5651***	-4.3103***	-0.2026
	(1.10)	(3.56)	(-0.08)	(-0.58)	(2.66)	(-7.53)	(-0.98)
Dollar trading volume	0.1384***	0.1513***	-0.0310**	-0.0037	0.1398***	-0.1650***	-0.1559***
	(15.10)	(16.54)	(-2.33)	(-0.26)	(10.14)	(-6.22)	(-11.43)
Amihud illiquidity	14.8887***	17.0413***	5.8548***	12.2765***	17.3478***	-23.0240***	7.6641***
	(11.37)	(13.49)	(3.88)	(7.96)	(11.33)	(-9.17)	(5.19)
Observations	20,597	20,597	20,597	20,597	20,597	20,580	20,500
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	38.77	38.77	38.77	38.77	38.77	38.78	38.81