Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

The Effects of Volatility on Liquidity in the Treasury Market

Andrew Meldrum, Oleg Sokolinskiy

2023-028

Please cite this paper as:
Meldrum, Andrew, and Oleg Sokolinskiy (2023). "The Effects of Volatility on Liquidity in the Treasury Market," Finance and Economics Discussion Series 2023-028. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2023.028.

The Effects of Volatility on Liquidity in the Treasury Market*

Andrew Meldrum[†]   Oleg Sokolinskiy[‡]

April 27, 2023

Abstract

We study the relationship between volatility and liquidity in the market for on-the-run Treasury securities using a novel framework for quantifying price impact. We show that at times of relatively low volatility, marginal trades that go with the flow of existing trades tend to have a smaller price impact than trades that go against the flow. However, this difference tends to diminish at times of high volatility, indicating that the perceived information content of going against the flow is less when volatility is high. We also show that market participants executing trades aggressively using market orders will experience larger increases in price impact than those executing trades passively using limit orders as volatility increases. And times of low market depth are associated with increased risk of high price impact and high sensitivity to volatility in future, perhaps because liquidity is more reliant on high-speed quote replenishment and is therefore more fragile.

Keywords: liquidity, Treasury market, market depth, volatility, order execution, hidden Markov model

JEL codes: G01, G10, G12, C51, C58

*This article represents the views of the authors, and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or other members of its staff. We are grateful for helpful comments from and discussions with Dobrislav Dobrev, Michael Fleming, Yesol Huh, and Edith Liu, as well as participants at a Federal Reserve System seminar.

[†]Board of Governors of the Federal Reserve System. Email: andrew.c.meldrum@frb.gov

[‡]Board of Governors of the Federal Reserve System. Email: oleg.v.sokolinskiy@frb.gov

1 Introduction

Understanding liquidity conditions in the market for on-the-run U.S. Treasury securities is important to market participants and policymakers alike because of the important role the market plays in the financial system. It is well known that liquidity in the Treasury market tends to be inversely related to interest rate volatility. At times of high volatility, the risks to market makers associated with adverse price movements are high, so market makers will tend to provide less liquidity and charge more for the liquidity they provide. In this paper, we deepen this existing understanding of how volatility and liquidity are related in three ways. First, we show that the strength of the perceived signal from trading against the existing trade flow for the future direction of prices tends to diminish as volatility increases. Second, we show that the extent to which market participants experience a deterioration in liquidity as volatility increases will depend on the strategies they use to execute trades. And third, we show that the sensitivity of liquidity to volatility tends to be higher when liquidity is lower, with the probability of entering a low-liquidity / high-sensitivity state depending on the willingness of market makers to post quotes in large volumes.

These conclusions are based on a model of the impact of trades and orders on prices. In an illiquid market, trades and orders have larger price impact. Our novel framework for estimating price impact allows us to consider several important dimensions of liquidity on the central limit order book (CLOB) platforms that account for a large share of trading in on-the-run Treasury securities. A CLOB consists of bid and ask queues of "limit orders" (that is, quotes) to buy or sell specified quantities of securities, respectively, with the best bid price being below the best ask price. An important aspect of CLOB-based trading that matters for the price impact of executing trades is that market participants have choices about how to execute. One option is to consume liquidity on the CLOB by submitting "market orders" to buy or sell specified quantities of a security at the best available quoted prices; equivalently, they can submit limit orders that "cross the spread," that is, with prices such that they can be immediately matched with other limit orders resting on the other side of the order book. We term market orders and limit orders that cross the spread "marketable orders." Conditional on liquidity providers having posted sufficient quotes to the other side of the order book, marketable orders will immediately become trades and therefore provide fast and guaranteed execution. A second option is to execute trades by *providing* liquidity, that is, by posting additional limit orders with prices such that they cannot be immediately matched, in the anticipation that the orders will be matched to marketable orders placed by other market participants later. Executing a trade using such "non-marketable" orders may be cheaper than using marketable orders because it avoids paying the bid-ask spread but comes at the risk that the order will not be matched quickly (or at all).[1] Another dimension of an execution strategy that matters for the price impact of a large "parent order" is how the market participant splits it into smaller "child orders" spread over time to lower the overall price impact.

To allow us to examine how these dimensions of execution strategies affect price impact, our model specification permits trade flow imbalances and non-marketable order flow imbalances to have different effects on prices and to affect prices non-linearly, unlike in the majority of previous studies. The regression-based approach was first established by Kyle (1985), although our specific implementation is close in spirit to Adrian et al. (2023), in that we estimate price impact separately on a day-by-day basis based on higher-frequency observations. Inclusion of non-marketable order flow imbalance is important because of previous evidence that order flow imbalance is superior to trade flow imbalance in explaining price changes (see, for example, Cont et al. (2014)).[2] Indeed, we confirm that a model that ignores non-marketable order flow imbalance would be mis-specified. We also show that it would be inappropriate to assume identical effects to marketable and non-marketable order flow imbalances, as is often done when order flow is considered in models of

[1] In practice, a sophisticated execution strategy may use a mixture of marketable and non-marketable orders (see, for example, Moro et al. (2009)).

[2] The results of Cont et al. (2014) were subsequently extended by Xu et al. (2018) and Cont et al. (2021) to incorporate order flow imbalance at multiple levels of the CLOB.

price impact. We further show that trade flow and net non-marketable order flow imbalances both tend to affect prices sublinearly, that is, the power coefficients on both flow imbalances are less than unity. Theoretical justification for sublinearity is provided by Farmer et al. (2013), who develop a game-theoretic model of an algorithmic execution service with agents endowed with varying information; with this framework, the price impact of parent orders scales sublinearly with size. Sublinearity is also consistent with previous empirical evidence (see Bouchaud et al. (2009) for a review). We show that the degree of sublinearity tends to diminish as volatility increases, which we argue is consistent with the perceived information content of orders that go against the existing flow being less at times of high volatility.

Armed with this model of price impact, we next turn to the question of how volatility affects the price impact of alternative execution strategies. We start with a baseline strategy under which a parent order is executed using market orders spread across the course of a single trading day using a "volume-weighted average price" (VWAP) strategy—that is, the parent order is spread in proportion to aggregate trade volume in each 1-minute time interval. This benchmark seems reasonable for various reasons. Pham et al. (2020) demonstrate that spreading large parent orders over time yields considerable savings in terms of price impact. Fraenkle et al. (2011) demonstrate the optimality of VWAP execution, while Frei and Westray (2015) argue that market participants consider it to be a fair execution benchmark, and Moro et al. (2009) provide evidence from equity markets that parent orders appear to be split across time in line with the overall market activity.[3] Using this baseline strategy, we reaffirm previous results that price impact is positively related to interest rate volatility (see, for example, Cont and Bouchaud (2000), Farmer et al. (2004) and Chordia et al. (2005a)).

By varying the details of the execution algorithm, we document several new findings about the importance of execution strategy and how it relates to interest rate volatility. Of particular note, price impact is generally greater for aggressive exe-

[3]While it is not straightforward to achieve VWAP execution because it requires an accurate forecast of the share of each intraday period in aggregate trading volume, Białkowski et al. (2008) suggest such a model.

cution using marketable orders than for passive orders using non-marketable limit orders, confirming that the previous findings of Hautsch and Huang (2012b) and Brogaard et al. (2019) for the equity market also apply to the Treasury market.[4] We show that the volatility of interest rates is a key factor driving the differences between the price impacts of aggressive and passive execution strategies: As volatility increases, the price impact of aggressive execution increases relatively quickly, likely because of the greater value of certainty about execution time in volatile markets. We also show that a relatively unsophisticated investor that spreads execution evenly over time rather than in proportion to aggregate volumes will experience relatively high price impact. Perhaps surprisingly, the benefits of more sophisticated execution tend to diminish at times of high interest rate volatility; a plausible explanation is that the relatively high information content of small order flows tends to diminish at times of high volatility. Finally, we show that the price impacts of alternative execution strategies are strongly positively correlated over time, suggesting that for many purposes we would not obtain misleading inferences about the overall dynamics of market liquidity by focusing on our baseline execution strategy.

We next examine how the level of price impact and its sensitivity to volatility relate to another popular liquidity measure, the total amount of quotes at the best prices resting on the order book, known as "market depth." In principle, market depth at the top of the book captures the volume of market orders that can be executed instantaneously without moving prices, so intuitively should be related to price impact. However, market depth has only limited value as a proxy for price impact because the large majority of willingness to post quotes is unobserved. As mentioned above, when a market participant wishes to execute a large parent order, rather than exhausting all available quotes at the best prices, they will typically split it into a number of smaller child orders, each of which are matched with a portion of the posted quotes, which are then quickly replenished (see, for example, Moro et al. (2009)). And because market participants adapt to changing market environments,

[4]Bikker et al. (2007) have also previously argued that there is a trade-off between the speed of execution and the associated price impact.

a deterioration in depth may not translate perfectly and immediately into higher price impact. Specifically, in response to lower depth, market participants may reduce the size of their child orders to avoid trading through multiple levels of the book.[5] In an analogous manner, if quotes at the best prices are temporarily exhausted, market participants tend to avoid executing at relatively unattractive prices and instead wait for quotes at better prices to be replenished, as shown by Dobrev and Meldrum (2020). These adaptions can mitigate the increase in price impact in the face of falling depth. While price impact and depth are negatively correlated, price impact tends to recover faster following an increase in volatility, suggesting that trading patterns quickly adapt to a lower level of depth.[6]

To better understand how the sensitivity of price impact to volatility varies over time, we estimate a three-state hidden Markov model (HMM) in which price impact is increasingly sensitive as we move from "good-liquidity" state, through a "medium-liquidity" state to a "bad-liquidity" state. Our approach is related to that of Flood et al. (2016), who link liquidity states identified using an HMM to observed economic variables. We extend their modeling framework by integrating explanatory variables directly into the HMM. Specifically, we allow the probabilities of transitioning between the states to depend on market depth. We show that while periods of low depth are not always associated with high price impact contemporaneously, low depth does appear to raise the probability of price impact being high and more sensitive to volatility—that is, low depth makes liquidity more fragile. This finding has an intuitive interpretation: the lower depth is, the more rapidly it must be replenished to support a given flow of market orders without moving prices substantially. If a shock makes the high-frequency trading firms who replenish quotes relatively quickly less willing to engage in that activity, or causes trading volumes

[5] Hautsch and Huang (2012a) and Hautsch and Huang (2012b), among others, show that orders potentially trading through multiple levels of the book tend to be shunned by market participants.

[6] Given our study of the relationship between price impact and market depth, our paper is related to those that examine the signals from multiple measures of liquidity. While one strand of the related literature has shown that various liquidity measures contain much common information (e.g. Chordia et al. (2000), Chordia et al. (2003)), Næs et al. (2008), and Korajczyk and Sadka (2008)), several other studies have focused more on the different signals from different liquidity measures, as we do.

that overwhelm the replenishment of quotes, then incoming order flow may start to have a greater impact on prices (see Board of Governors of the Federal Reserve System (2020)).

Our study is related to Hautsch and Huang (2012a), who show a link between price impact and depth in that the price impact of a limit order depends on how deep in the CLOB it rests. The results of Nguyen et al. (2020) and Aronovich et al. (2021) similarly point to a rich dependence structure between liquidity, volume, and depth. Nguyen et al. (2020) develop a model for the interaction of liquidity and volatility at high-frequencies; they find that market depth and trading volume significantly affect intraday volatility. Our study complements these previous findings at a trading-day time scale, in that our HMM estimates show that lower market depth leads to states characterized by higher price impacts of trade flows and, thus, all else equal, greater volatility. Other relevant previous studies include Chollete et al. (2007), who argue that there can be major disagreements between liquidity measures, especially during stress events. In addition, Fraenkle et al. (2011) allow the slope and intercept in price impact regressions to depend on explanatory variables such as trading volume and volatility, which highlights how other liquidity measures can directly affect price impact, but they do not consider market depth.

The remainder of this paper proceeds as follows. In Section 2, we set out our model for estimating price impact, discuss evidence of nonlinearities in trade and non-marketable order flow imbalances, and explain how these nonlinearities relate to interest rate volatility. In Section 3, we explore how price impact depends on execution strategy and how this relates to interest rate volatility. In Section 4, we analyze how the level of price impact and its sensitivity to volatility depend on market depth. In Section 5, we offer some concluding remarks.

2 Sublinearity of Price Impact

2.1 Model of Price Impact

Our measure of price impact is based on a model relating returns to net trade and non-marketable order flows. The model takes the form:

$$\begin{aligned} r_{t-1,t} = {} & \beta_m sign\left(TF_{t-1,t}\right)\left|TF_{t-1,t}\right|^{\gamma_m} \\ & + \beta_n sign\left(NMOF_{t-1,t}\right)\left|NMOF_{t-1,t}\right|^{\gamma_n} + \varepsilon_t. \end{aligned} \tag{1}$$

where $t = 0, \ldots, T$ are the boundaries of one-minute intraday periods, which together form a partition of the studied trading day; $TF_{t-1,t}$ is the trade flow imbalance, measured as the difference between buyer- and seller-initiated trade volumes between $t-1$ and t; $NMOF_{t-1,t}$ is the non-marketable order flow imbalance, measured as order flow imbalance as calculated in Cont et al. (2014) between $t-1$ and t, less the effect of trades, which are accounted for in trade flow;[7] and $r_{t-1,t} = \frac{P_t - P_{t-1}}{P_{t-1}}$ is the simple return based on changes in mid-quoted prices between $t-1$ and t (P_{t-1} and P_t, respectively). The inclusion of both trade flow and non-marketable order flow as explanatory variables, with different impacts on prices, is motivated by the results of Brogaard et al. (2019), who show that in equity markets market orders individually have greater price impact than limit orders, but that limit orders contribute more to price discovery overall. The feature that the effects of trade flow and non-marketable order flow are nonlinear is motivated by the review in Bouchaud et al. (2009), and, in particular, the asymmetric liquidity hypothesis of Lillo and Farmer (2004) and Farmer et al. (2006). These studies reconcile the fact that the direction of order flow imbalances displays long memory with market efficiency: an order has a smaller price impact when following another order in the same direction.

We estimate the model in equation (1) for the 10-year on-the-run Treasury Note

[7]Specifically, $NMOF_{t-1,t} = OFI_{t-1,t} - TF_{t-1,t}$, where $OFI_{t-1,t} = \sum_{n=N(t-1)+1}^{N(t)} \left[q_n^b I\left(P_n^b \geq P_{n-1}^b\right) - q_{n-1}^b I\left(P_n^b \leq P_{n-1}^b\right) - q_n^s I\left(P_n^s \leq P_{n-1}^s\right) + q_{n-1}^s I\left(P_n^s \geq P_{n-1}^s\right)\right]$; $N(t)$ is the index of the last order book event—a market order, limit order, or cancellation—of the period between $t-1$ and t; q_n^b and q_n^s are the sizes of the queue at the best bid price, P_n^b, and ask price, P_n^s, respectively, at event n; and $I(.)$ is an indicator function taking the value 1 when the argument is true and 0 otherwise.

using nonlinear least squares separately for each day between March 1, 2014 and March 31, 2023. We omit days with shortened trading hours, leaving a total of 2,215 days. We consider only the most active trading hours, between 7 a.m. and 4.45 p.m., giving a maximum of 585 one-minute observations per trading day. For model estimation purposes we retain only intraday periods with at least one trade. We also constrain the values of γ_m to be below 1.8, which avoids extreme estimates of this parameter on a small proportion (less than 1 percent) of days.

Figure 1 shows histograms of the daily parameter estimates. The upper-left and lower-left panels show estimates of the slope coefficients on trade flow and non-marketable order flow, that is, β_m and β_n, respectively. The estimates of β_m are generally larger than those of β_n, which is consistent with trade flow having a larger price impact than the same non-marketable order flow. The upper-right and lower-right panels show estimates of the nonlinearity coefficients γ_m and γ_n, respectively. Both coefficients are less than one on the large majority of trading days (83 percent and 84 percent, respectively), implying that price impact is sublinear. Based on likelihood ratio tests, we reject the null of linearity at the 5 percent significance level on 66 percent of days.

2.2 Understanding Sublinearity

The price impact of a marginal trade flow x_t and non-marketable order flow y_t between times $t-1$ and t is given by

$$\theta\left(x_{t-1,t}, y_{t-1,t}\right)_{t-1,t} = \left|NPI\left(x_{t-1,t}, y_{t-1,t}\right) - NPI\left(0,0\right)\right|, \tag{2}$$

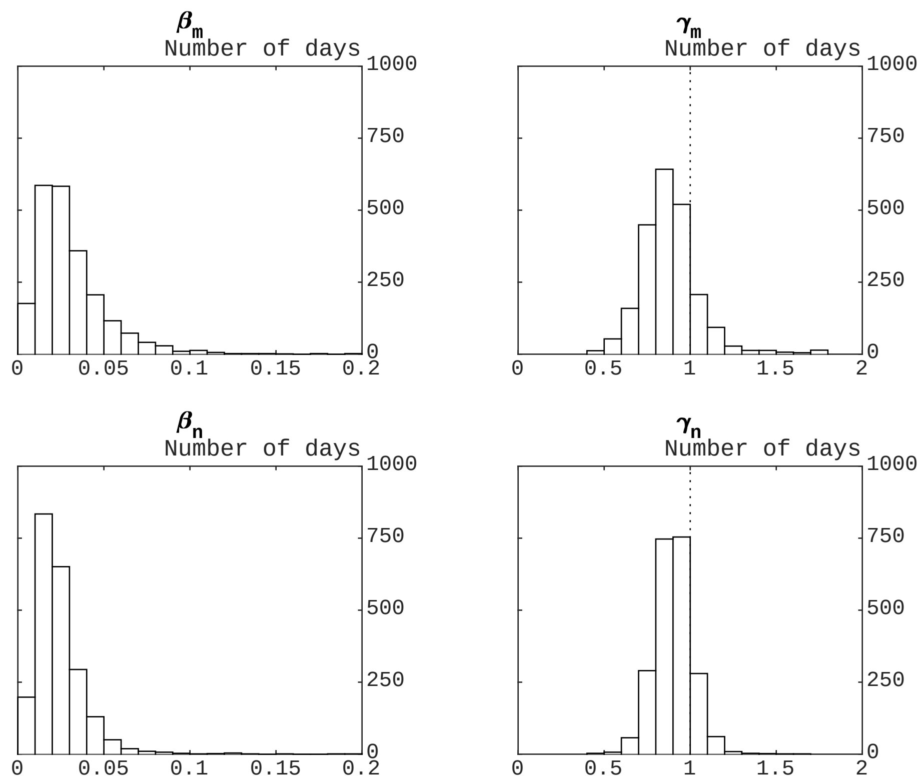
where

$$\begin{aligned} NPI\left(x_{t-1,t}, y_{t-1,t}\right) = {} & \beta_m sign\left(TF_{t-1,t} + x_{t-1,t}\right)\left|TF_{t-1,t} + x_{t-1,t}\right|^{\gamma_m} \\ & + \beta_n sign\left(NMOF_{t-1,t} + y_{t-1,t}\right)\left|NMOF_{t-1,t} + y_{t-1,t}\right|^{\gamma_n}; \end{aligned} \tag{3}$$

is the conditional mean of the simple return given the existing $TF_{t-1,t}$ and $NMOF_{t-1,t}$, and the contributions to flows from an additional $x_{t-1,t}$ in marketable orders and

Figure 1: Estimates of Nonlinearity Parameters for Price Impact

The panels show histograms of the daily estimates of the parameters from equation (1) between March 1, 2014 and March 31, 2023. Histograms for β_m and β_n, in the top and bottom left panels, respectively, use bin widths of 0.01. Histograms for γ_m and γ_n, in the top and bottom right panels, respectively, use bin widths of 0.1.



Sources: Repo Inter Dealer Broker Community; authors' calculations.

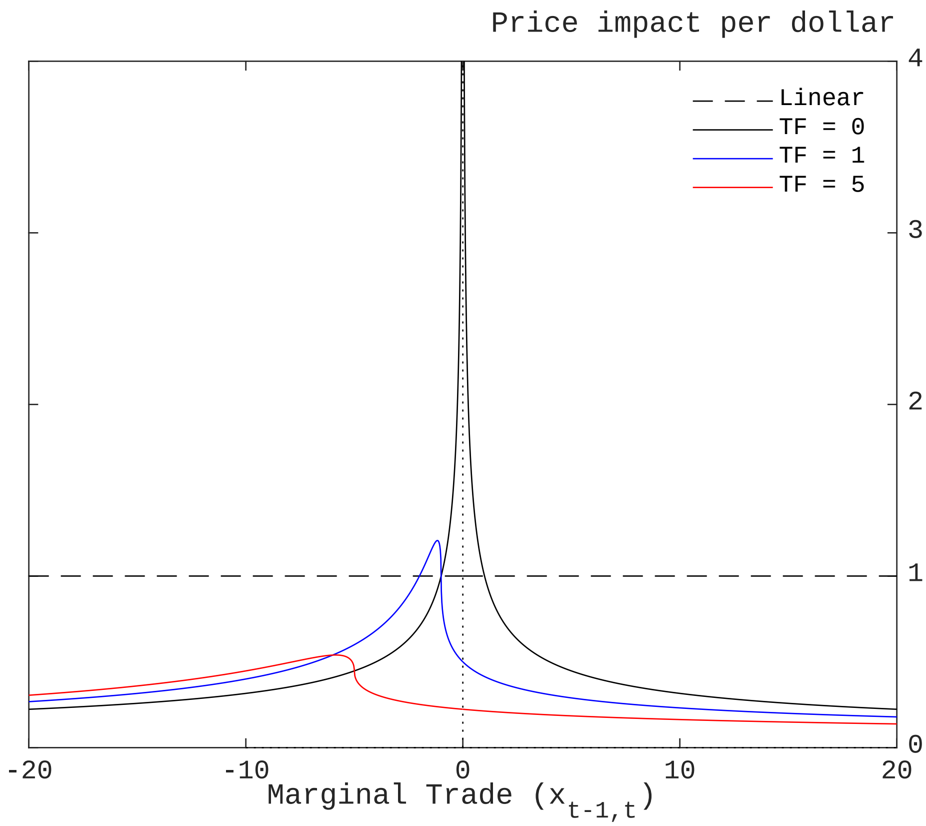
$y_{t-1,t}$ in non-marketable orders.

Figure 2 shows a stylized depiction of the relationship between the marginal trade flow, $x_{t-1,t}$, on the horizontal axis and its per-dollar price impact, $\frac{\theta(x_{t-1,t},0)_{t-1,t}}{|x_{t-1,t}|}$, on the vertical axis, for various values of existing trade flow imbalance $TF_{t-1,t}$ and with β_m set equal to one. The case of a linear model—that is, with $\gamma_m = 1$—is shown by the dashed line; in this case, the per-dollar contribution to price impact does not depend on either the existing trade flow imbalance or the size of the marginal trade. The various solid lines show the cases of a sublinear model—with $\gamma_m = 0.5$—for $TF_{t-1,t}$ equal to 0 (black line), 1 (blue line), and 5 (red line). Consider first the case with $TF_{t-1,t} = 0$. In this case, a small marginal trade has a greater per-dollar impact on prices than a large marginal trade. One intuition for this property is that relatively sophisticated investors are seen as more likely to split their trades into relatively small chunks (see, for example, Barclay and Warner (1993)). Small order flows therefore likely represent an increase in the activity of relatively informed market participants, which contain a stronger signal for future prices, and hence a proportionally higher price impact relative to the size of the trade. Conversely, large order flows have a proportionately smaller price impact relative to the size of the trade because they are more likely to be made by comparatively uninformed market participants.

When the existing trade flow is non-zero, the plotted curve changes materially. First, the curve is no longer symmetric. A marginal trade that goes in the opposite direction to the existing trade flow imbalance has a larger per-dollar price impact than a marginal trade that goes in the same direction as the existing trade flow imbalance. The intuition is that a trade that goes against the current flow of the market is perceived as being more likely to reflect some new information about the direction of prices than one that goes with the flow. The largest per-dollar price impact is for a marginal trade that exactly offsets the existing trade flow; the reason that the per-dollar price impact starts to decrease as the magnitude of the marginal trade increases beyond that point is that the marginal trade has the effect of flipping the sign of the overall trade flow imbalance.

Figure 2: Illustration of Nonlinearity in Trade Flow Imbalance

The figure provides a stylized illustration of how the per-dollar contribution of a marginal trade varies with the size of the marginal trade. The horizontal axis represents $x_{t-1,t}$ and the vertical axis represents $\frac{\theta\left(x_{t-1,t},0\right)_{t-1,t}}{\left|x_{t-1,t}\right|}$. The dashed line represents a linear model with $\gamma_m = 1$ and the solid lines represent a sublinear model with $\gamma_m = 0.5$, for three levels of $TF_{t-1,t}$. For this illustration we set $\beta_m = 1$.



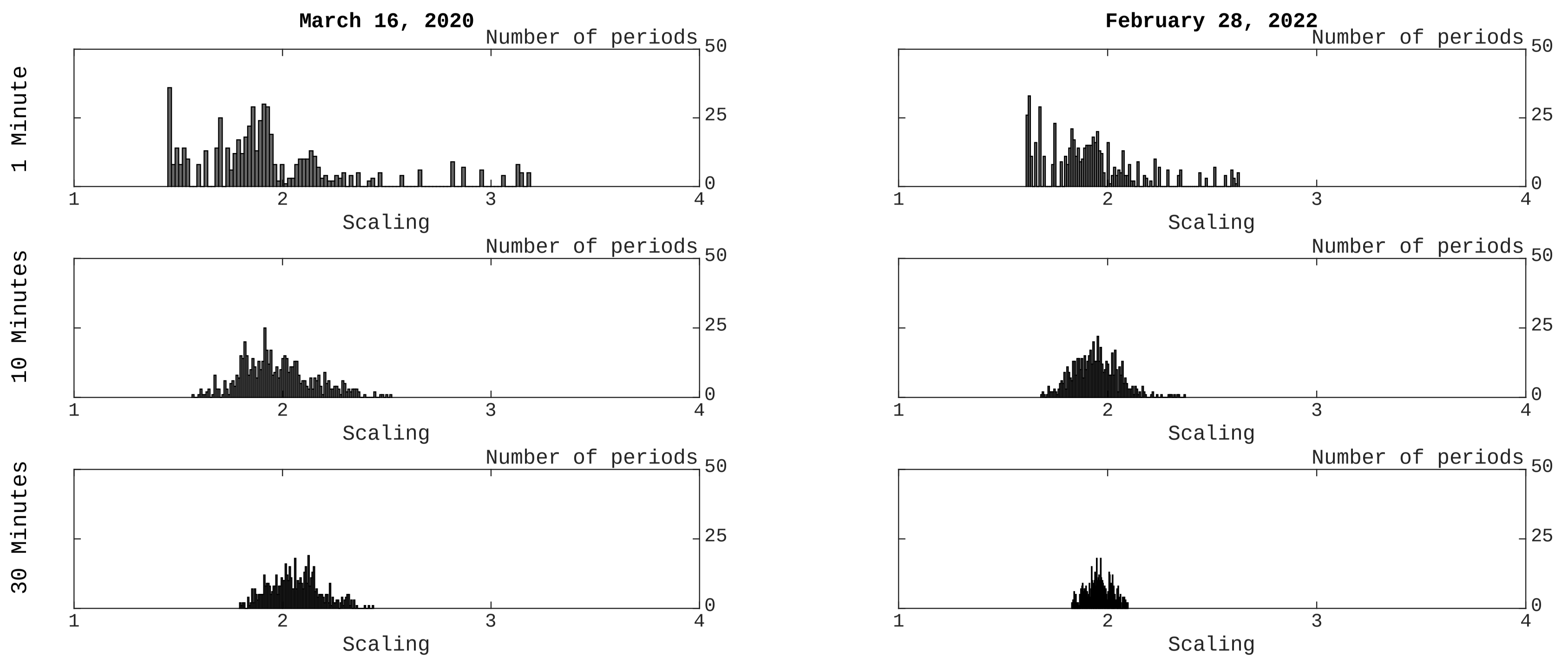
Source: Authors' calculations.

This feature of our model implies that the scaling of price impact with the size of marginal trade and order flows is subject to uncertainty, in that it depends on the imbalances of the market overall at the time the trade is executed or the order is submitted. To illustrate this uncertainty, we compute the ratio between a trade of $20 million and one of $10 million executed using market orders over a 1-minute period. The first row of charts in Figure 3 show histograms of these ratios for all possible 1-minute periods on two days: one, March 16, 2020, at the height of the COVID-pandemic-related stresses, and the other, February 28, 2022, a more typical day. The estimates of γ_m on these two days were 0.62 and 0.75, respectively. As a result of this sublinearity, in the majority of 1-minute periods of these days the scaling was sublinear, with ratios below 2. However, there were notable proportions of 1-minute period with ratios above 2, implying that market participants who cannot forecast aggregate order flows accurately face greater costs of scaling up to executing larger orders than those who can; a sophisticated market participant will attempt to execute a marginal trade at times when order flow is already imbalanced in the direction of the marginal trade. We will return to the remaining panels of Figure 3 in Section 3.1.

While sublinearity encourages fast execution of large volumes if a market participant wants to trade in the same direction as the net of other participants' flows, we cannot extrapolate the finding of sublinearity to an unlimited extent, beyond the point where flows would be inconsistent with functional market dynamics. Market participants' ability to execute large orders quickly and with diminishing price impact is constrained by the amount of quotes available at the best prices and the willingness of market makers to replenish quotes quickly in response to incoming orders. While trading through multiple levels of the CLOB may occur, it is a rare event that is both costly to the market participant and may scare off liquidity providers, hampering execution for some time. Market participants therefore operate predominantly in the region of trade flows that would not cause such localized market dysfunction.

Figure 3: Scaling with Parent Order Size

The figure shows histograms of price impact of a parent order divided by the price impact of a second parent order with size half that of the first, executed at all possible times within the trading day. The left column shows results for March 16, 2020 and the right column shows results for February 28, 2022. The upper, center, and lower rows show results for execution of first parent orders of size $20 million, $200 million, and $600 million over horizons of 1 minute, 10 minutes, and 30 minutes, respectively. All trades are executed targeting volume-weighted average price. In each case, the estimates are split into 100 equal-width bins.



Sources: Repo Inter Dealer Broker Community; authors' calculations.

2.3 Sublinearity and Interest Rate Volatility

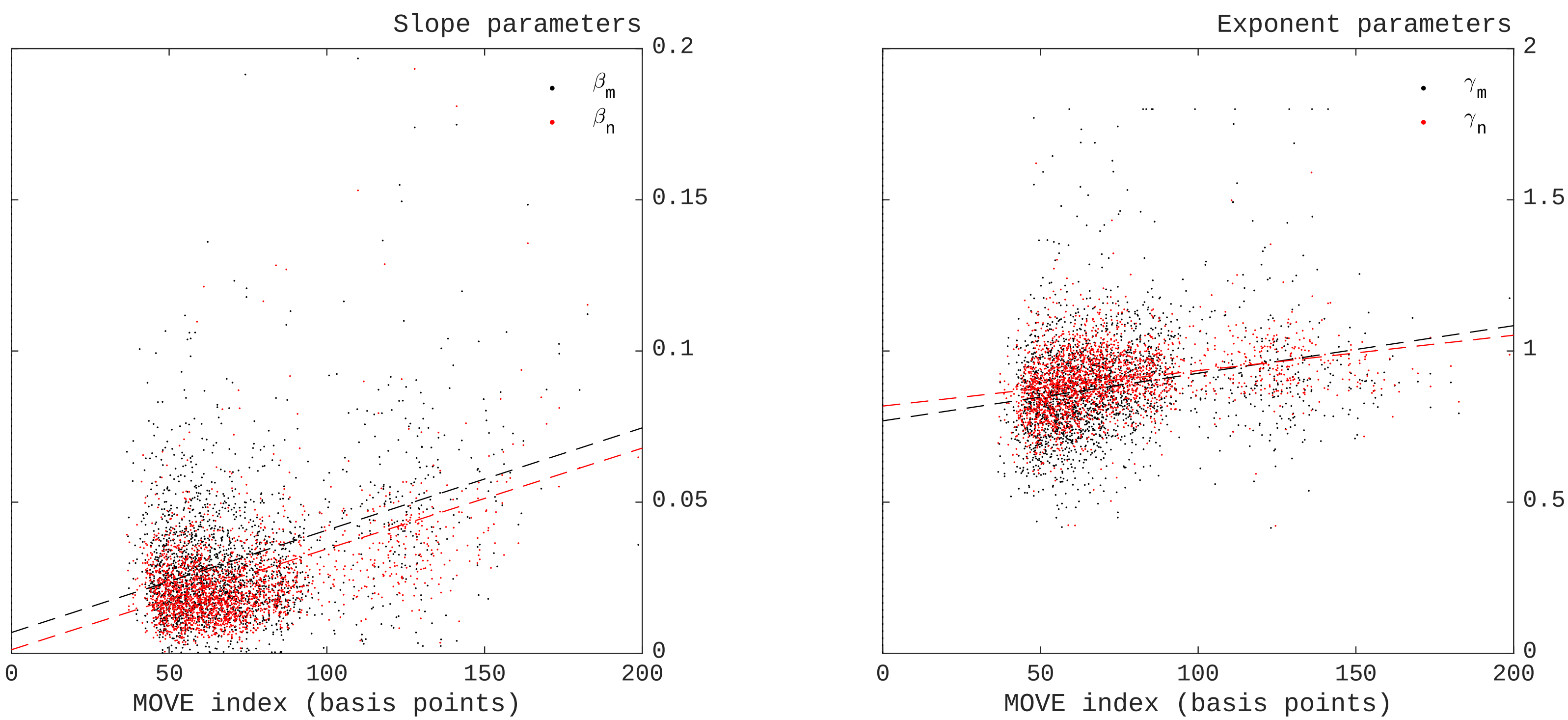
We next turn to the relationship between the parameter estimates and interest rate volatility. Figure 4 plots the daily estimates of the slope parameters from equation (1) in the left panel, and the exponent parameters in the right panel. In both panels, the parameters are plotted against the MOVE index, which is a weighted average of the volatilities implied by one-month options on Treasury securities with various maturities, and which we source from Bloomberg. Each marker corresponds to estimates for a single day. The dashed lines show fitted values from linear regressions of the parameters on a constant and the MOVE index, estimated using ordinary least squares (OLS). We highlight two key points: First, the estimated slope coefficients in the left panel tend to increase with volatility, with the slopes of the regression lines being positive and significantly different from zero at the 1 percent significance level. This result is consistent with standard findings that liquidity conditions are inversely related to volatility. Second, as shown in the right panel, the impact of trade and non-marketable order flows on prices tends to shift from being generally sublinear on average toward being closer to linear on average as volatility increases. The slope coefficients implied by the dashed regression lines are again positive and significantly different from zero at the 1 percent level. Conditional on the MOVE being above 100 basis points, the mean values of γ_m and γ_n are 0.95 and 0.96, respectively. This finding that price impact becomes close to linear at relatively high levels of volatility suggests that the perceived proportionally high information content of relatively small order flow tends to diminish on high-volatility days. That could be because during volatile periods fundamental economic drivers are overwhelmed by noise trading due to portfolio re-balancing by investors seeking to control their risk exposures.

3 Price Impact and Execution Strategy

In a framework that allows trade and non-marketable order imbalances to have different and nonlinear effects on prices, there is not a single measure of price impact

Figure 4: Relationship between Price Impact Regression Parameters and Volatility

The figure shows estimates of the parameters β_m and β_n in the left panel and γ_m and γ_n in the right panel, as obtained from estimating equation (1) using daily data for the period between March 1, 2014 and March 31, 2023. Each marker represents a single day; a small number of markers fall outside the plotted ranges. The estimates are plotted against the MOVE index of the Treasury market implied volatility. The dashed lines show linear regression lines estimated using data for all days.



Sources: Repo Inter Dealer Broker Community; Bloomberg Finance LP, Bloomberg Per Security Data License; authors' calculations.

for the execution of a large parent order —it depends more realistically on the time horizon for execution, how the parent order is split over time into child orders, and to what extent it is executed aggressively using marketable orders or passively using non-marketable orders. In this section, we explore how these dimensions affect estimates of price impact. In Section 3.1, we explain how we estimate price impact for parent orders: We set out our baseline strategy targeting VWAP and pursuing aggressive execution over a trading day. In Section 3.2, we show that our baseline estimates of price impact tend to increase with volatility, which extends previous findings to show the nonlinearity in the relationship between volatility and price impact. In Section 3.3, we analyze the effects of varying the execution strategy and how these depend on volatility.

3.1 Baseline Execution Strategy

We estimate the price impact of a parent order of size ξ dollars spread over day d, $\theta_d(\xi)$, as the cumulative effect on returns over the day from the additional marginal stream of trade flow $\{x_t\}_{t=1}^{T}$ and non-marketable order flow $\{y_t\}_{t=1}^{T}$, relative to the scenario where no additional flows occurred.[8] Formally, this is given by

$$\theta_d(\xi) = \left| \prod_{t=1}^{T} \left(1 + NPI\left(x_{t-1,t}, y_{t-1,t}\right) - NPI\left(0,0\right)\right) - 1 \right|. \tag{4}$$

Our baseline strategy considers the price impact of a purchase of $500 million—equivalent to about 1.5 percent of median daily trade volumes over the period—executed using marketable orders spread across one-minute periods through the day, in proportion to the volume traded within each minute. In this baseline case, we have $x_{t-1,t} = \xi \frac{V_{t-1,t}}{V_d}$ and $y_{t-1,t} = 0$, where $V_{t-1,t}$ and V_d are the existing trad-

[8] We recognize that some portion of price impact may prove to be transient: a transient price impact framework described in Bouchaud et al. (2009) is an alternative to nonlinear price impact in the sense that it may also reconcile long memory of trade flows with market efficiency. Such transience may play a lesser role in our framework as we already allow for price impact to be nonlinear in trade flow, with the impact of each new trade dependent on some history of past market events. We assume that such a transient component of price impact (if any) is relatively small and decays sufficiently slowly for its effect to be of secondary importance for calculations over the timescales adopted in this paper.

ing volumes between times $t-1$ and t, and in aggregate for day d, respectively.[9] This splitting in child orders corresponds to an ideal VWAP algorithm, named so because it targets the volume-weighted average price over its execution horizon. While a market participant would not in practice be able to forecast trading volumes perfectly, as discussed above, we can nevertheless think of the resulting measure of price impact as being representative of the impact of aggregate trading, as well as corresponding to a recognized execution benchmark.

Of course the choice of the \$500 million parent order over the period of one trading day is somewhat arbitrary, in that we could have picked a range of sizes that are not outsized relative to aggregate daily volume. As discussed in Section 2.2, a feature of a nonlinear model is that the scaling of price impact with trade size may be different on each day. Thus, there may be a concern that the results we discuss below are not robust to varying the trade size. However, it turns out that the scaling becomes close to linear as the execution horizon lengthens to cover the active hours of the entire trading day. This result means that our main conclusions are likely to be robust to varying the size of the parent order. To demonstrate this point, the second and third rows of Figure 3 show histograms of the same ratios as the first row, which were discussed above, but for parent orders executed over all possible 10-minute periods (the second row) and 30-minute periods (the third row). To keep the average amount executed in each minute constant, the experiments for execution over 10 minutes and 30 minutes consider parent orders of \$200 million and \$100 million, and of \$600 million and \$300 million, respectively. In all cases, execution is using market orders following VWAP targeting. As the execution horizon lengthens the ratios converge toward 2—even on these days with estimates of γ_m well below one, implying that market participants can become increasingly confident that scaling will be approximately linear as the execution horizon lengthens. It appears that periods in which scaling is sublinear and superlinear roughly balance out, leaving the overall scaling roughly linear. This result is consistent with those of Bouchaud

[9] To simplify calculations, we do not account for the contribution of the marginal trades to volumes, which is a reasonable approximation due to their relatively small size.

et al. (2009), who demonstrate how the compensating effect of correlated trades results in approximately linear price impact when it is aggregated over a sufficiently large number of individual transactions.

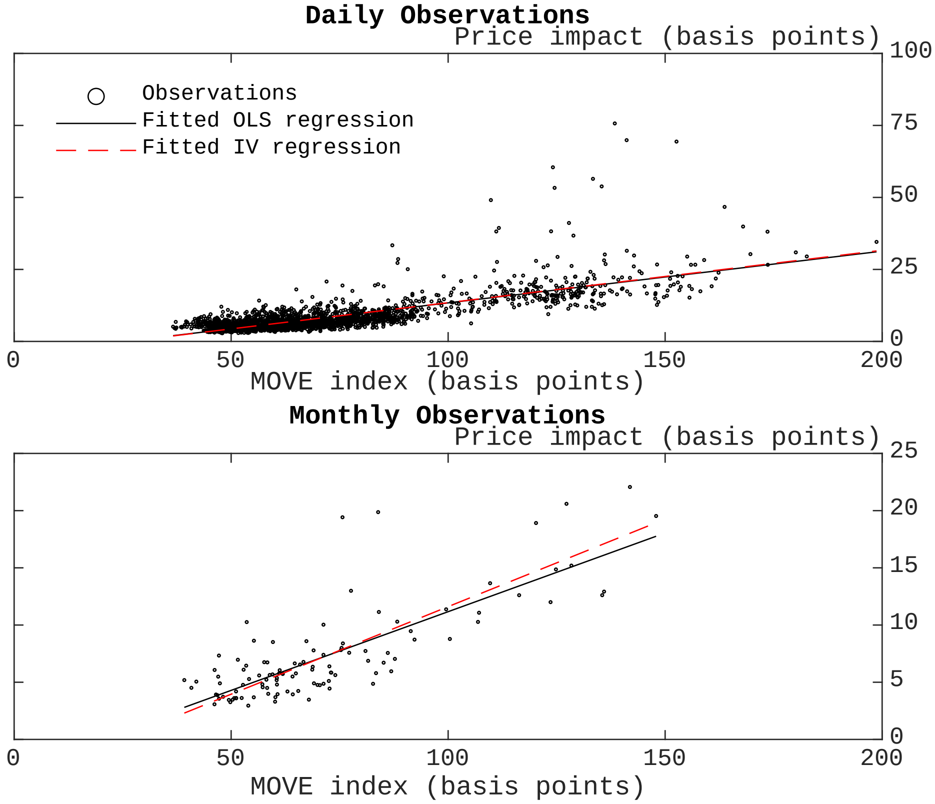
3.2 Price Impact and Volatility

The upper panel of Figure 5 plots our resulting baseline measure of price impact on the vertical axis against interest rate volatility on the horizontal axis. Each marker corresponds to a single trading day. We measure interest rate volatility using end-of-day observations of the MOVE index. Each black circle represents a single day in the sample. A fitted linear relationship estimated by OLS is shown by the black line. The slope is positive and statistically significantly greater than zero at the 5 percent significance level. This result that liquidity conditions are worse at times of high volatility confirms previous findings by Chordia et al. (2005a), among others; intuitively, market intermediaries are less willing to provide liquidity when the risk of large price moves is relatively high.

A potential concern with this analysis is that price impact and volatility are jointly determined. As volatility increases, liquidity conditions deteriorate, as captured by an increase in price impact. However, the increase in the price impact of a given trade may in turn result in prices becoming more volatile, resulting in a feedback loop; such intuition is consistent with previous findings by Chordia et al. (2005b) and Chordia et al. (2005a) that measures of volatility and liquidity show two-way Granger causality. Thus, the OLS estimate of the relationship between volatility and price impact may be subject to an endogeneity bias. In an attempt to guard against the potential endogeneity bias, Figure 5 also shows a (red dashed) regression line estimated using the first lag of volatility to instrument for volatility. The instrumental variables (IV) estimator of the slope is essentially the same as the OLS estimate when using daily data and remains statistically significant at the 5 percent level. While lagging volatility by a single day may not be adequate to guard against endogeneity bias due to persistence in volatility, when using monthly averages of daily data, shown in the bottom panel, the IV estimate of the slope

Figure 5: Price Impact and Volatility

The figure shows scatter plots of the MOVE index of Treasury market volatility (horizontal axes) against estimated price impact (vertical axes). Each marker represents a single day between March 1, 2014 and March 31, 2023. The upper panel shows daily estimates and the lower panel shows monthly averages. The lines report fitted values from a bivariate linear regression with a constant, estimated by OLS (black and solid line) or IV (red and dashed line)



Sources: Repo Inter Dealer Broker Community; Bloomberg Finance LP, Bloomberg Per Security Data License; authors' calculations.

remains close to the OLS estimate.

Another potential concern is that a linear relationship between volatility and price impact may be mis-specified. Indeed, inspection of the scatter plots in Figure 5 suggests that price impact is more sensitive to volatility at higher levels of volatility than at lower levels of volatility. In Section 4, we explicitly model this nonlinearity in the volatility–price impact relationship.

3.3 Varying the Execution Strategy

3.3.1 Aggressive versus Passive Execution

Aggressive execution using marketable orders has the advantage of greater certainty about the schedule and price of execution, allowing for completion over a shorter time-frame, thereby reducing the risk that prices move due to news or activity of other market participants between the time when the decision to trade is made and the time when the trade is executed. But this lower risk comes at the cost of crossing the spread and a potentially higher price impact due to marketable orders being perceived by other market participants to contain greater information. Passive execution, on the other hand, waits for the market to come to the execution algorithm, thereby saving on execution costs. But it gives up some control over the scheduling and, thus, also brings risks of prices shifting before orders are filled.

In practice, execution strategies may employ a mix of aggressive and passive execution, with a market participant's risk tolerance and trading strategies determining the specific mix, and the strategy may be amended as the execution progresses (for example, the proportion of aggressive execution may be increased if execution falls behind a desired schedule). For simplicity, we compare the price impact of the extreme cases of 100 percent aggressive and 100 percent passive execution. Specifically, we compare results for aggressive versus passive VWAP execution of \$500 million spread over one trading day. Formally, passive execution price impact is computed using $x_{t-1,t} = 0$ and $y_{t-1,t} = \xi \frac{V_{t-1,t}}{V_d}$ in equation (3). This estimate corresponds to an ideal passive execution strategy where the algorithm manages to post child orders at the best quotes and there is sufficient demand for posted liquidity to enable immediate execution. It does not account for the possibility that the price moves away from posted order price and the subsequent response of the algorithm. For example, suppose the initial bid-ask spread is from 99-01 to 99-02, the algorithm posts a sell order at 99-02, and the market consequently moves to a bid-ask spread of 99-00 to 99-01—that is, away from the sell order. Suppose also that the algorithm responds by canceling the original sell order at 99-02, which is now at the second level of the order book, and places a replacement order at the new top of the book, that is, at at 99-01. Our approach implicitly assumes that the cancellation and new quote would have exactly offsetting impacts on prices.

Table 1 reports the correlation between the time series of price impacts corresponding to various execution strategies. The correlation between price impact using aggressive and passive VWAP execution (models A and B, respectively) is high, at 0.979. That said, there is a fairly systematic difference in the levels of the estimates under aggressive and passive execution, which is illustrated by Figure 6. The left panel shows a histogram of the difference in basis points between the price impacts under aggressive and passive execution. On the large majority of days, aggressive execution has greater price impact to compensate for the uncertainty of the execution schedule associated with passive execution, as shown by the fact that the large majority of the mass of the histogram falls to the right of zero. To shed some light on what drives the difference between aggressive and passive execution, the right panel shows how interest rate volatility measured by the MOVE index is associated with a greater difference. At relatively low levels of volatility, there is a relatively weak relationship between volatility and the price impact differential of aggressive relative to passive execution on that day. However, at relatively high values of volatility, there are more days in which the price impact of aggressive execution notably exceeds that of passive execution. We venture two possible explanations: First, greater immediacy and certainty delivered by aggressive execution is particularly valuable in more volatile markets. And second, volatility may better hide the information content of non-marketable flow, while the impact of trades is

Table 1: Correlations between Estimates of Price Impact

The table reports correlations between price impacts from following various execution strategies for the parent order of selling $500 million over a trading day. The sample cover the period from March 1, 2014 to March 31, 2023. The column headed "Execution" refers to whether execution is by VWAP or TWAP, and "Agg./Pass." refers to whether execution is aggressive (A) or passive (P).

	Execution	Agg./Pass.	Correlation			
			A	B	C	D
A	VWAP	A	1	0.979	0.972	0.979
B	VWAP	P	0.979	1	0.962	0.997
C	TWAP	A	0.973	0.967	1	0.973
D	TWAP	P	0.979	0.997	0.973	1

Sources: Repo Inter Dealer Broker Community; authors' calculations.

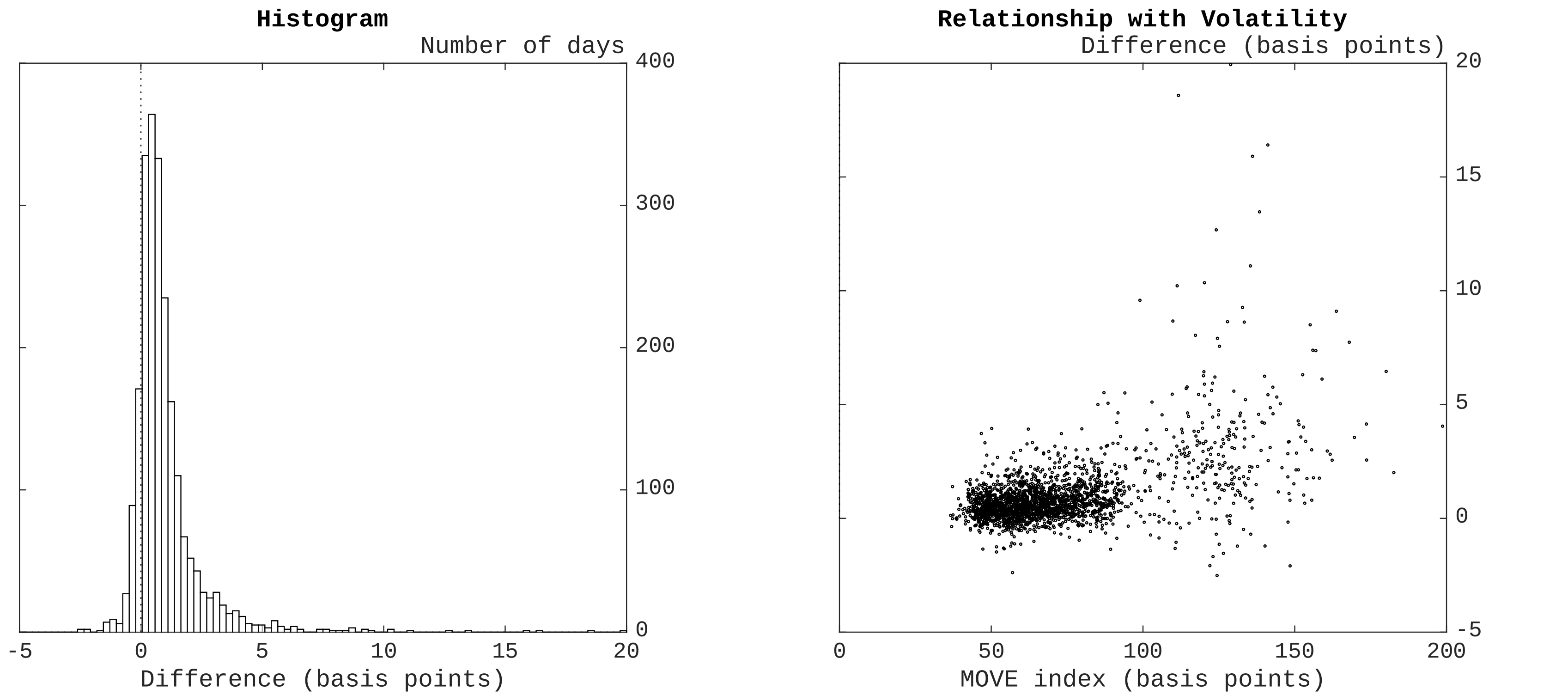
more mechanical.

3.3.2 Splitting Order Execution over Time

In this section, we consider whether price impact is different for a relatively unsophisticated investor that is unable to execute trades using something approximating a VWAP strategy. As an illustration, we consider a TWAP strategy, which does not require forecasting of aggregate trade or order flow because the amount that is executed in each time interval is simply $\frac{\xi}{T}$; in this section, we assume that execution is 100 percent aggressive. Table 1 reports that price impacts from TWAP execution (models C and D in the table) are highly positively correlated with those under VWAP execution (models A and B), with the lowest correlation being above 0.96. The left panel of Figure 7 shows a histogram of the difference between price impact under TWAP and VWAP execution. The difference is positive on the large majority, 83 percent, of days, meaning that VWAP execution generally achieves a lower price impact, as we would expect. However, VWAP execution does not always out-perform TWAP execution, as there is a small number of days on which the latter achieves a lower price impact. The right panel shows a scatter plot of the difference between price impact under VWAP and TWAP plotted against the MOVE volatility index. It shows that the advantage of VWAP generally appears to lessen as volatility increases. This is perhaps surprising, in that it implies that the more sophisticated execution strategy actually has less of an advantage at times

Figure 6: Passive versus Aggressive Execution

The left panel shows a histogram of price impact under aggressive execution minus the price impact of passive execution over the period March 1, 2014 to March 31, 2023. The estimates are split into 100 equal-width bins. The right panel plots this difference against the MOVE index of Treasury market volatility, where each marker represents a single day.



Sources: Repo Inter Dealer Broker Community; Bloomberg Finance LP, Bloomberg Per Security Data License; authors' calculations.

of market stress. The reason is that on days with very low volatility there are a relatively large number of 1-minute period with zero trades. As discussed above, when net trade flow is zero, the price impact of a marginal trade is at its highest. Unreported results show that if we drop periods with zero trades from consideration, TWAP does not perform as badly on low-volatility days.

4 Price Impact, Sensitivity to Volatility, and Market Depth

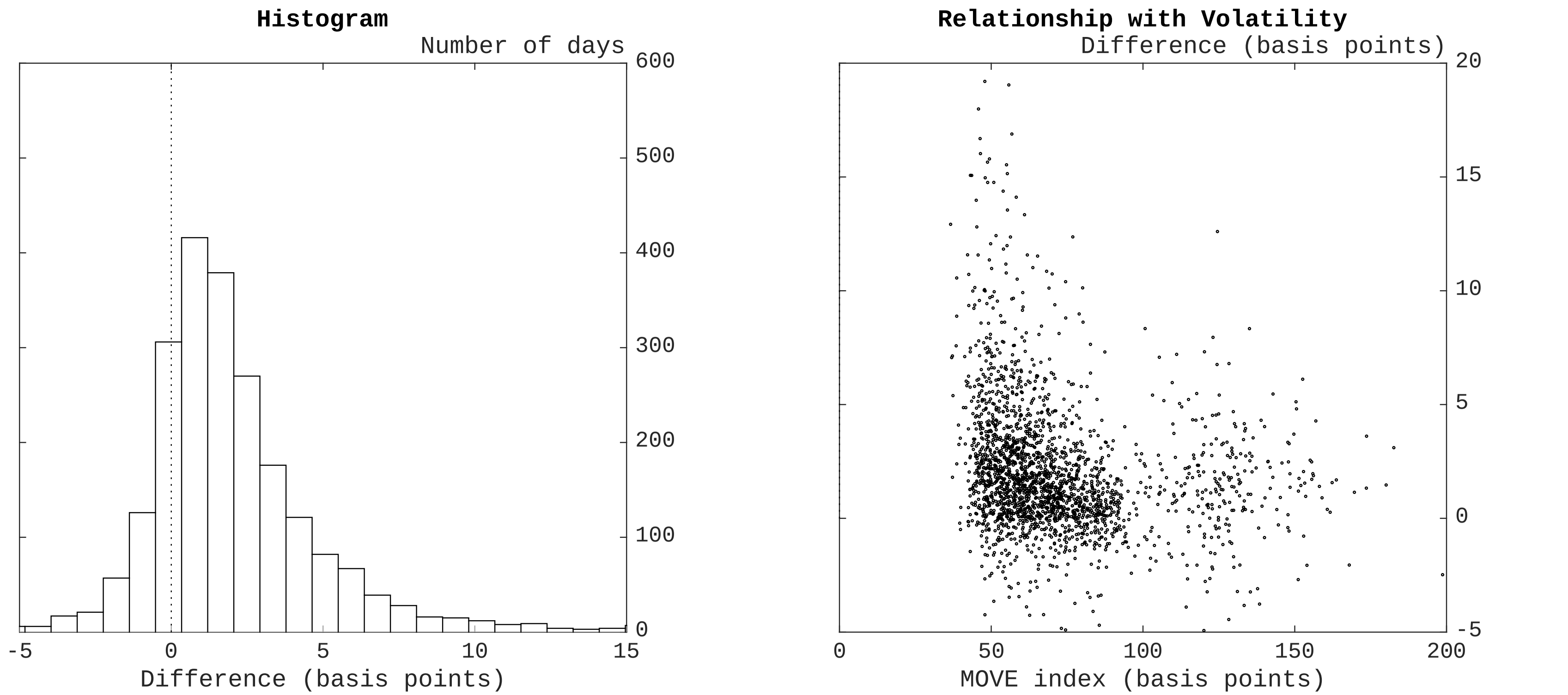
In this section, we examine how price impact and its sensitivity to volatility varies with market depth. In Section 4.1, we show that our baseline estimate price impact is negatively correlated with market depth contemporaneously but that price impact can have notably different time-series dynamics than market depth. In Section 4.2, we present a Hidden Markov Model (HMM) relating price impact to depth; the results of the model suggest that times of low depth are associated with a greater risk of increases in price impact and its sensitivity to volatility in future.

4.1 Price Impact and Market Depth

Figure 8 shows daily time series of price impact and market depth. We compute depth as the time-weighted mean of the total posted quantities at the best bid and ask prices between 7 a.m. and 4.45 p.m. The upper two panels show the same price impact estimates but using different vertical axis scales, while the lower panel shows market depth. The sample correlation between price impact and market depth is fairly negative, at -0.75, which is intuitive because falling depth and rising price impact are both associated with a deterioration in liquidity conditions. The period following the onset of the COVID-19 pandemic in March 2020 stands out as seeing by far the highest values of price impact, as well as the largest decline in depth. There were also notable spikes in price impact coinciding with other days of local troughs in depth identified by Aronovich et al. (2021), as indicated by the vertical dashed lines: the Treasury market “flash crash” of October 15, 2014, where prices moved

Figure 7: Algorithm for Splitting a Parent Order into Child Orders

The left panel shows a histogram of price impact under TWAP minus the price impact under VWAP execution, for the period March 1, 2014 to March 31, 2023. The estimates are split into 100 equal-width bins. The right panel plots this difference against the MOVE index of Treasury market volatility. A small number of data points fall outside the axis limits of both panels.



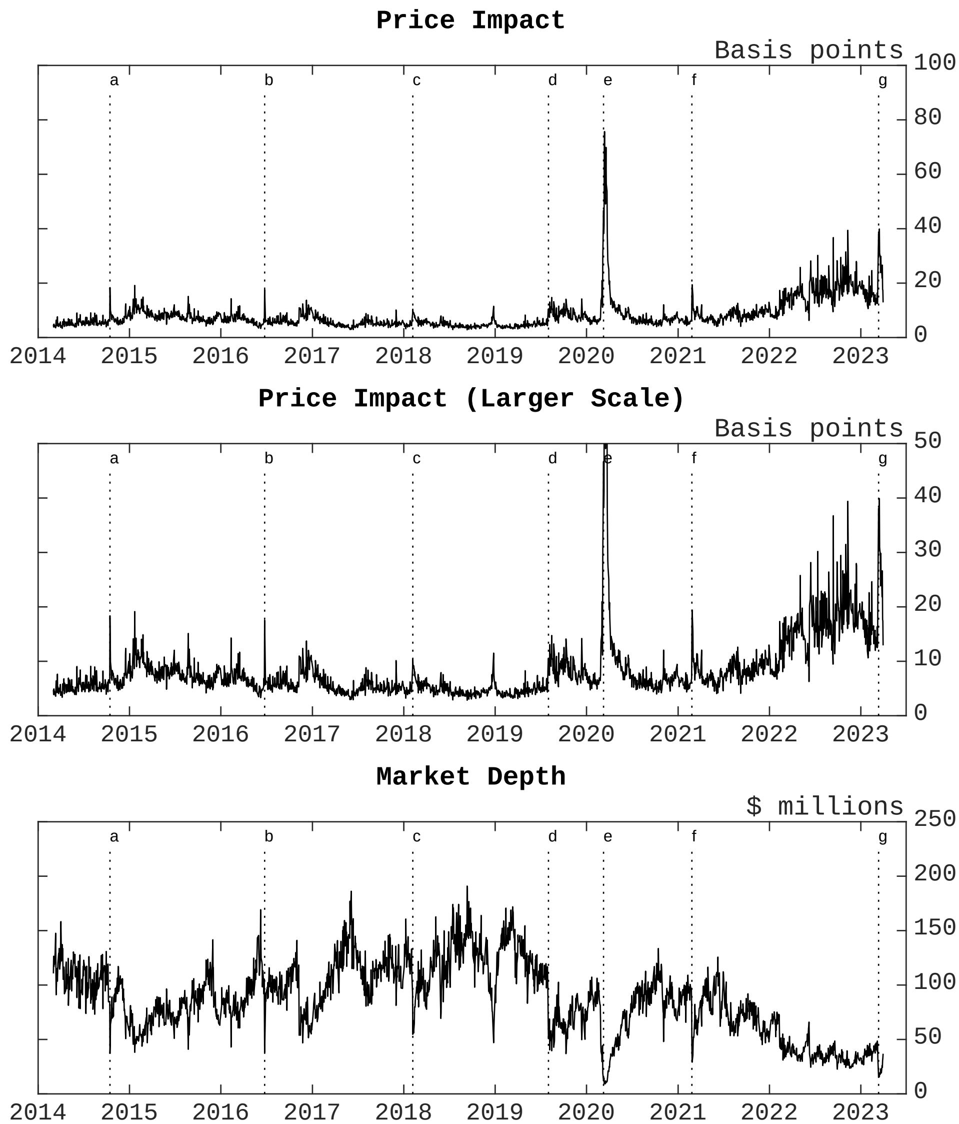
Sources: Repo Inter Dealer Broker Community; Bloomberg Finance LP, Bloomberg Per Security Data License; authors' calculations.

sharply before reversing quickly; the U.K. "Brexit" referendum on June 24, 2016; the sharp spike in the VIX index of equity market volatility on February 6, 2018; the onset of concerns among financial market investors about the global growth outlook on August 2, 2019; a sharp worsening in the market strains associated with the COVID-19 pandemic on March 9, 2021; and the flash event of February 25, 2021. In addition, we pick out the Monday following the closure of Silicon Valley Bank, that is, March 13, 2023. There were also less sharp increases in price impact and declines in depth around early 2015 and late 2016, which were likely associated with spillovers from stresses in European sovereign debt markets and the notable increase in yields following U.S. elections, respectively. And depth has declined and price impact has risen notably since late 2021, which market commentary has linked to high interest rate volatility associated with uncertainty about the economic outlook.

Although peaks in price impact tend to coincide with troughs in depth, price impact and depth appear to have substantially different time-series dynamics. For example, the sample autocorrelation functions in Figure 9 show that depth is substantially more persistent than price impact. We also observe substantial differences between the dynamics of market depth and price impact following the onset of episodes of market stress. Figure 10 shows depth and price impact from 10 days before to 15 days after the seven events highlighted in Figure 8. To facilitate comparison across episodes, we normalize all variables to have the value of unity at the beginning of the event window. Depth recovered relative quickly following the Brexit referendum on June 24, 2016, and the VIX spike on February 6, 2018, although it still took 2 or 3 weeks before depth had retraced the large majority of the initial decline. In comparison, the recoveries following the flash crash of October 15, 2014, the August 2, 2019 global growth concerns, the February 25, 2021 flash event, and onset of the March 2023 banking-sector stresses were slower. And the recovery following the most severe episode of market dysfunction, the onset of the COVID-19 pandemic in March 2020 took substantially longer still. Price impact rose immediately at the onset of all episodes. In some episodes, price impact fell back relatively rapidly compared with the slower recovery of depth; this observation

Figure 8: Price Impact and Market Depth

The charts show times series of price impact (upper and center panels) and market depth (lower panel) for the 10-year Treasury Note. The upper and center panels plot the same estimates but using different y-axis scales. The vertical dashed lines correspond to: (a) October 15, 2014; (b) June 24, 2016; (c) February 6, 2018; (d) August 2, 2019; (e) March 9, 2020; (f) February 25, 2021; and (g) March 10, 2023.

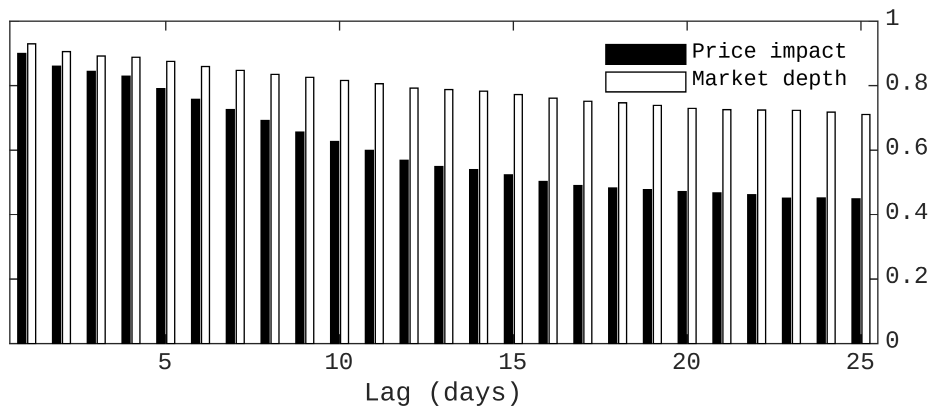


Sources: Repo Inter Dealer Broker Community; authors' calculations.

is consistent with Aronovich et al. (2021), who show that bid-ask spreads tend to recover substantially faster than market depth following these episodes. A plausible explanation for this is that market participants rapidly adjusted to lower depth by trading in smaller sizes to allow quotes on the order book to be replenished. In other cases—most notably March 2020 and March 2023—the peak of price impact did not come until several days after the trough in depth, suggesting that the low level of depth may increase the risk of a spike in price impact.

Figure 9: Autocorrelation Functions of Price Impact and Market Depth

The bars show the sample autocorrelations of market depth and our baseline estimate of price impact from March 1, 2014 to March 31, 2023 at various lags.

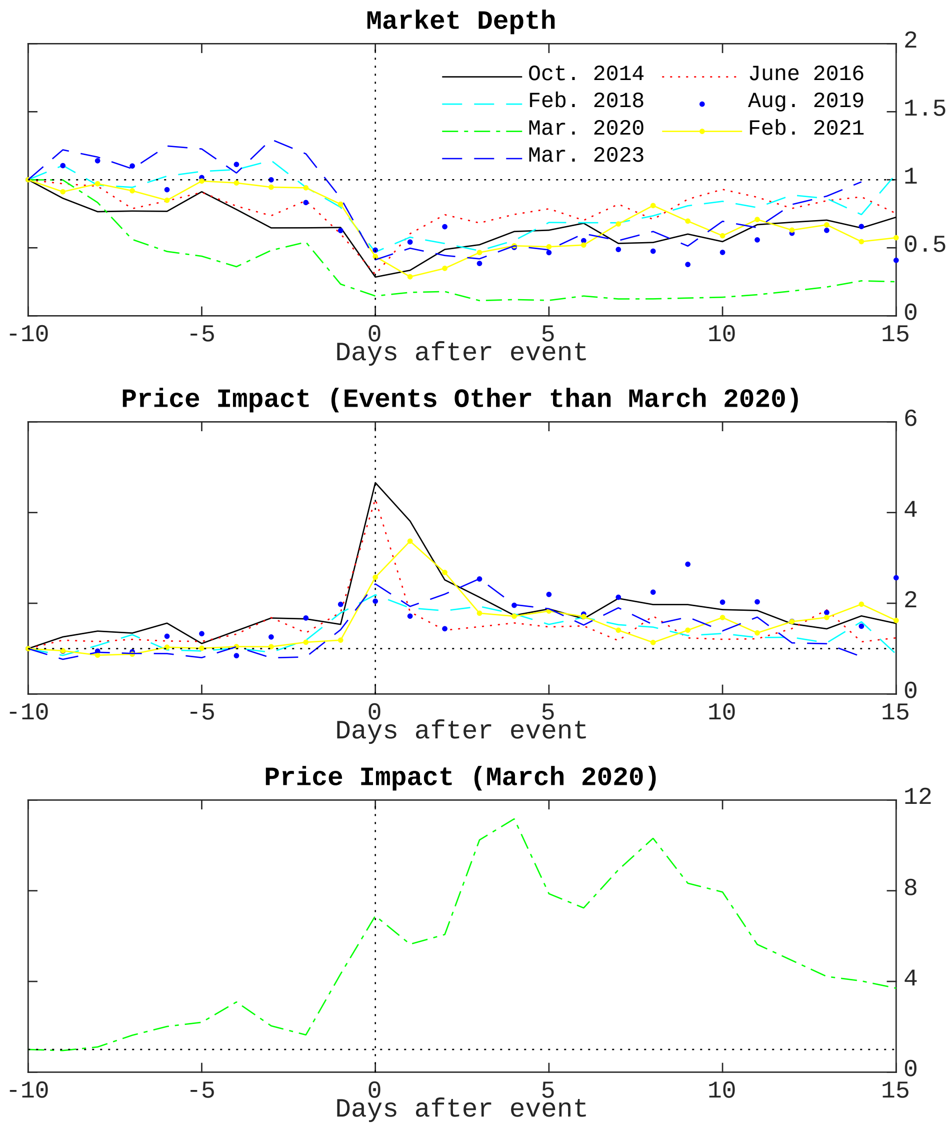


Sources: Repo Inter Dealer Broker Community; authors' calculations.

These results raise questions of how market depth and price impact are related, and whether it matters that depth remains low after a stress event even though price impact has returned to more normal levels. Given market participants' unwillingness to trade through multiple levels of the book, as noted in Hautsch and Huang (2012a), a decrease in top-of-the-book depth causes optimal execution to prescribe splitting the parent order into smaller child orders. Hypothetically, market makers may replenish the order book sufficiently rapidly to result in either no or moderate increase in the price impact of execution; the point that prompt order-book replenishment matters for the cost of trading has been made previously by Dobrev and Meldrum (2020) and Aronovich et al. (2021). The immediate reduction in depth following a market stress event is often associated with a pull-back in high-speed liquidity provision, but this liquidity provision returns faster than the willingness

Figure 10: Price Impact and Market Depth following Episodes of Known Liquidity Strains

The charts show market depth and price impact from 10 days before until 15 days after various episodes of known liquidity strains. All series are normalized to have value one 10 days before the event.



Sources: Repo Inter Dealer Broker Community; authors' calculations.

of market participants to quote in large size; hence, depth remains low but price impact falls back.

One potential interpretation of these results is that depth is not a meaningful indicator of liquidity conditions, because high-speed liquidity replenishment means that trading costs can be kept low even though most of the willingness of market participants' to provide liquidity remains hidden. However, a conjecture broadly in line with Raman et al. (2014) and Board of Governors of the Federal Reserve System (2020) is that greater dependence on sophisticated quoting and execution algorithms in such circumstances may nevertheless make liquidity more fragile, meaning that low depth makes it more likely that price impact could increase again if there is a further pull-back in high-speed liquidity provision—perhaps because high-frequency trading firms may be inclined to scale back their activity if some other stress event hits or because elevated trading volumes will overwhelm the speed of quote replenishment. Indeed, as mentioned above, some of the spikes in price impact following the events discussed above did not come until after sometime following the initial trough in depth. We therefore next study more systematically whether the low level of depth is associated with liquidity becoming more fragile, meaning an elevated risk of price impact rising.

4.2 Hidden Markov Model

Our analysis is based on an HMM. The model relates price impact to interest rate volatility according to

$$\theta_d = \beta_i MOVE_d + \epsilon_{i,d}, \tag{5}$$

where $MOVE_t$ is the MOVE index of interest rate volatility; and $\epsilon_{i,d} \sim \mathcal{N}(0, \sigma_i^2)$ for three latent liquidity states $i = 1, 2, 3$. This specification is motivated by the analysis in Section 3, which demonstrates that price impact is positively related to interest rate volatility and that price impact may become more sensitive to volatility as volatility increases. We label the states such that the slope coefficients decrease with the state index in the measurement equation (5); thus, state 1 is a "low-liquidity" state in which price impact is most sensitive to volatility, state 2 is a "medium-liquidity" state, and state 3 is a "high-liquidity" state.[10]

We consider two versions of the HMM. The first, Model I, has constant probabilities of transitioning between the states. The parameterization of the model ensures that the conditional probabilities of transitioning to the 3 states add to one. The second, Model II, further allows market depth to affect the transition probabilities. In this case, we model the transition probabilities using the multinomial logistic specification of Visser and Speekenbrink (2010) and Zucchini and MacDonald (2009). Specifically, let s_d denote the liquidity state on day d. Then, the time-varying probability of transitioning between states is

$$p_{i,j,d} = \frac{\exp\left(\delta_{i,j} + \gamma_{i,j} z_d\right)}{1 + \sum_{k=1,3} \exp\left(\delta_{i,k} + \gamma_{i,k} z_d\right)}, \tag{6}$$

where $p_{i,j,d}$, with $j \neq 2$, denotes the day-d probability that $s_{d+1} = j$ conditional on $s_d = i$, and z_d is market depth on day d. To ensure that probabilities add up (over j) to one, the probabilities of transitions to "medium-liquidity" state 2 must be parameterized as

$$p_{i,2,d} = \frac{1}{1 + \sum_{k=1,3} \exp\left(\delta_{i,k} + \gamma_{i,k} z_d\right)}. \tag{7}$$

As a consequence of this normalization, there are no parameter estimates for $\delta_{i,2}$ and $\gamma_{i,2}$. We estimate the model by maximum likelihood, using the functionality provided by Visser and Speekenbrink (2010).

4.3 Results from the HMM

Table 2 reports parameter estimates of the two HMMs. Considering first the measurement equation of the models reported in Panel $\mathcal{A}$, the main results are robust across specifications. The slope coefficients on the MOVE index are all positive and highly significant, meaning that higher volatility is associated with greater price

[10]Bayesian and Akaike information criteria favor models with three states to models with two states.

impact in all states. Moreover, the differences between MOVE coefficients among states are also large and significant, in line with the hypothesis that price impact is more sensitive to volatility when liquidity is low. In other words, changes in volatility by itself, without variation in the sensitivity of price impact to volatility, are insufficient to adequately explain the variation in price impact.

We next turn to the transition equations of the models, reported in Panel $\mathcal{B}$. Consider first Model I. If today's state is state 1 (the low-liquidity state), there is a probability of about three fourths of remaining in that state tomorrow, a probability of about one fourth of transitioning to state 2 (the medium-liquidity state), and negligible probability of transitioning to state 3 (the high-liquidity state). If today's state is the medium- or high-liquidity states, there are higher probabilities of remaining in the same state tomorrow, and only small probabilities of changing state.

Now consider Model II, in which market depth affects the transition probabilities. The estimates of $\gamma_{1,1}$, $\gamma_{2,1}$, and $\gamma_{3,1}$ are all negative and statistically significant at the 5 percent level. This result means that as depth decreases the probability of transitioning to the low-liquidity state 1 increases, which makes intuitive sense. It is less clear-cut whether an increase in depth increases the probability of transitioning to the high-liquidity state. The estimates of $\gamma_{1,3}$ and $\gamma_{3,3}$ are both positive, meaning that as depth increases the probability of transitioning to the high-liquidity state 3 increases. That said, the estimate of $\gamma_{1,3}$ is not statistically significant, possibly because of the number of transitions direct from the low- to high- liquidity states is small, so there is little information in the data to estimate this parameter precisely. The fact that the estimate of $\gamma_{2,3}$ is negative is less intuitive, although it is also not statistically significant.

To further illustrate the effect of depth on the transition probabilities, Figure 11 shows the transition probabilities conditional on various levels of depth, as implied by Model II. As depth decreases, the probability of transitioning to state 1 (the low-liquidity state) increases. For example, when depth is $10 million, the probability of transitioning to state 1 is above 0.8 regardless of the current state; and, in contrast,

Table 2: HMM Parameter Estimates

The table reports parameter estimates for HMMs using data on price impact, volatility, and, where relevant, market depth for the period from March 1, 2014 to March 31, 2023. Model I includes no explanatory variables in the transition equation and Model II includes market depth. Numbers in brackets indicate standard errors based on the Hessian. A *, **, or *** indicates two-sided p-values of less than 0.1, 0.05, and 0.10, respectively.

	Model I		Model II
	$\mathcal{A}$: Measurement Equation		
β_1	0.215***	β_1	0.210***
	(0.015)		(0.001)
β_2	0.130***	β_2	0.131***
	(0.001)		(0.001)
β_3	0.089***	β_3	0.089***
	(0.001)		(0.001)
σ_1	12.176***	σ_1	11.507***
	(1.037)		(0.845)
σ_2	1.916***	σ_2	1.711***
	(0.082)		(0.062)
σ_3	1.012***	σ_3	1.002***
	(0.025)		(0.025)
	$\mathcal{B}$: Transition Equation		
$p_{1,1}$	0.734***	$\delta_{1,1}$	11.554**
	(0.068)		(4.931)
$p_{1,1}$	0.258	$\delta_{1,3}$	-82.429
	-		(200.763)
$p_{1,3}$	0.007***	$\gamma_{1,1}$	-0.413**
	(0.012)		(0.179)
$p_{2,1}$	0.027***	$\gamma_{1,3}$	1.405
	(0.007)		(3.408)
$p_{2,2}$	0.930	$\delta_{2,1}$	3.892***
	-		(1.186)
$p_{2,3}$	0.043***	$\delta_{2,3}$	-1.634***
	(0.009)		(0.558)
$p_{3,1}$	0.002	$\gamma_{2,1}$	-0.144***
	(0.028)		(0.031)
$p_{3,2}$	0.028	$\gamma_{2,3}$	-0.013
	-		(0.008)
$p_{3,3}$	0.970***	$\delta_{3,1}$	2.075
	(0.006)		(2.269)
		$\delta_{3,3}$	-3.176***
			(0.936)
		$\gamma_{3,1}$	-0.057
			(0.038)
		$\gamma_{3,3}$	0.070***
			(0.012)

Sources: Repo Inter Dealer Broker Community; Bloomberg Finance LP, Bloomberg Per Security Data License; authors' calculations.

when depth is $120 million, these probabilities decrease to essentially zero.

Figure 12 shows the estimated smoothed probabilities of being in each latent liquidity state at each point in time, as implied by Model II, and computed using an algorithm from Visser and Speekenbrink (2010). The lower panel plots scaled price impact. The estimated state probabilities make intuitive sense, with higher probabilities of being in a low-liquidity state generally being higher at times of spikes in price impact. Until around late 2019 or early 2020, the probability of being in the high-liquidity state 3 was generally very high, with just a handful of episodes of being in the medium-liquidity state. These episodes generally coincided with moderate increases in price impact. Since 2020, however, the medium-liquidity state has shifted to being the predominant state, with the high-liquidity state becoming relatively rare and the low-liquidity state becoming much more common. The two most notable periods of low-liquidity in the sample appear to be associated with the COVID-related market turmoil in early 2020 and the low liquidity as economic uncertainty rose toward the end of the sample.

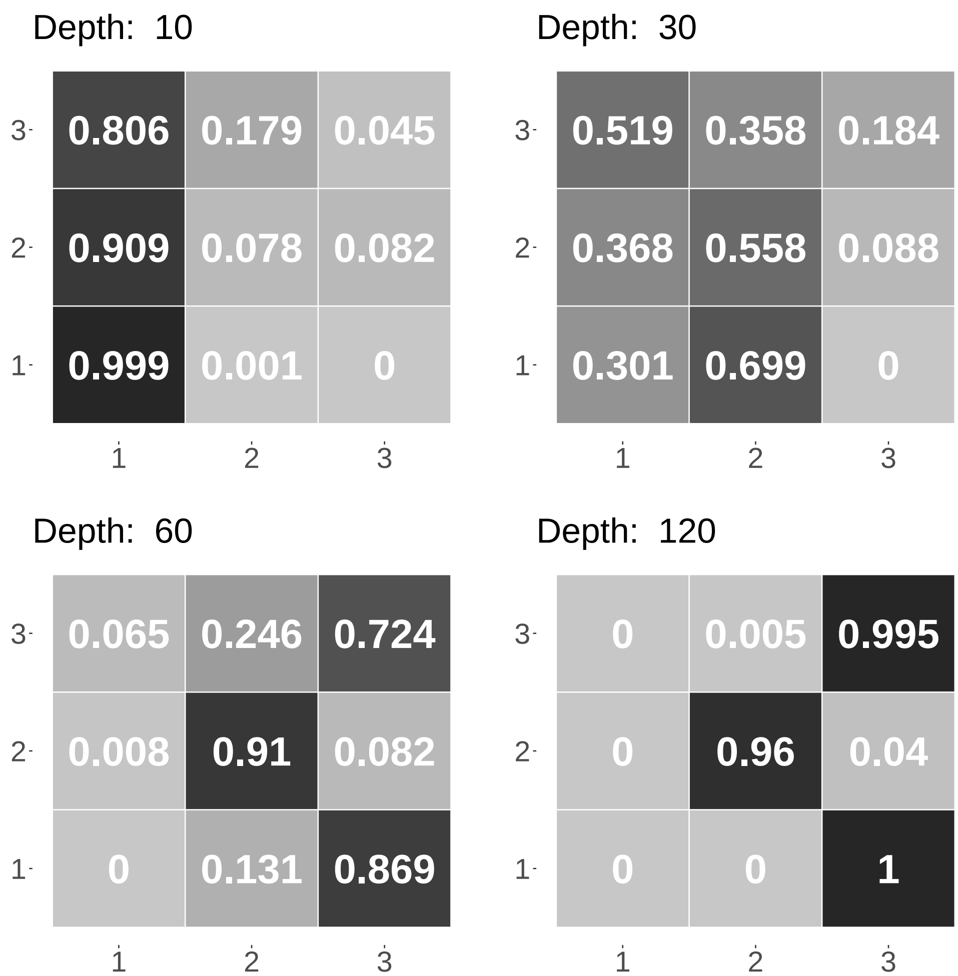
5 Conclusion

In this paper, we studied the relationship between volatility and liquidity in the market for on-the-run Treasury securities. Our analysis was based on a novel framework for quantifying price impact that allows trade flow imbalance and non-marketable order flow imbalance to have different and nonlinear effects on prices. We showed that the effect of flows on prices is generally sublinear, implying that marginal trades that go with the existing flow will have smaller price impact than trades that go against the existing flow. However, this sublinearity tends to diminish at times of high volatility, suggesting that the perceived relatively high information content of going against the flow lessens at times of high volatility.

We next examined the price impact felt by market participants seeking to execute a large trade over the course of a trading day. This price impact is negatively related to volatility, confirming previous results. However, we show that market participants

Figure 11: HMM Transition Matrix

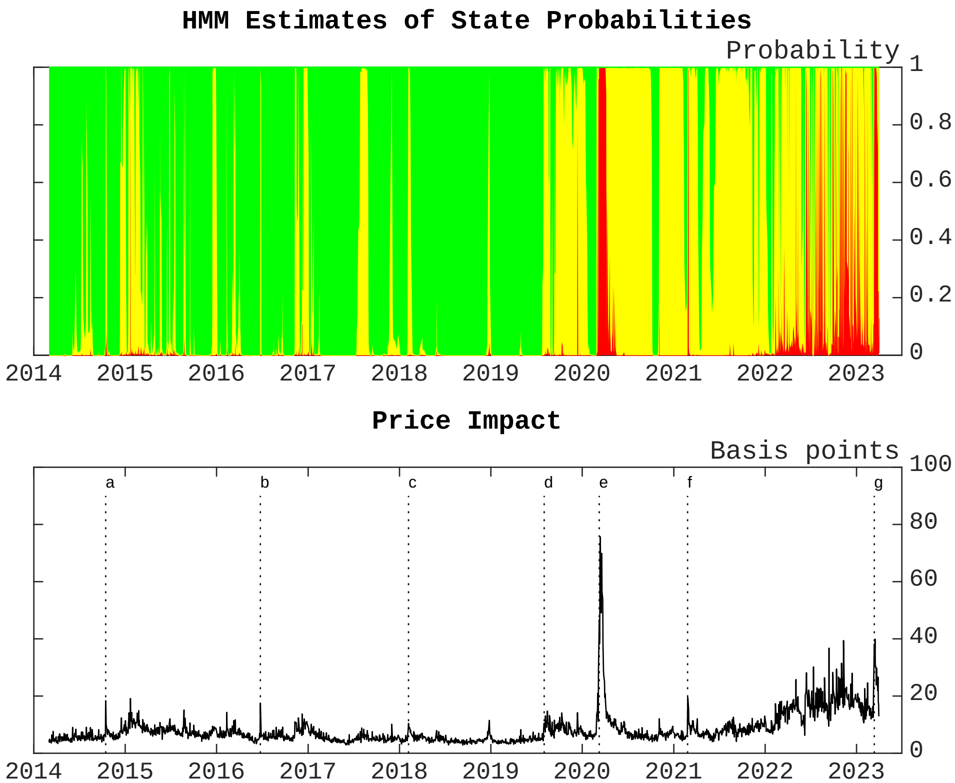
The figure shows the transition probability matrices from the HMM conditional on various levels of market depth, as implied by Model II. The levels of market depth are \$10 million, \$30 million, \$60 million, and \$120 million. In each case, the matrix shows the probabilities of transitioning from the states on the vertical axis to the states on the horizontal axis. Darker shading highlights higher probabilities.



Sources: Repo Inter Dealer Broker Community; Bloomberg Finance LP, Bloomberg Per Security Data License; authors' calculations.

Figure 12: HMM-Implied Probabilities of Being in Each State

The upper panel shows the estimated probabilities of being in each state, as implied by Model II. The green, yellow, and red areas indicate the probabilities of being in the high-, medium, and low-liquidity states, respectively. The lower panel shows a time series of price impact. The vertical dashed lines in the lower panel correspond to: (a) October 15, 2014; (b) June 24, 2016; (c) February 6, 2018; (d) August 2, 2019; (e) March 9, 2020; (f) February 25, 2021; and (g) March 13, 2023.



Sources: Repo Inter Dealer Broker Community; Bloomberg Finance LP, Bloomberg Per Security Data License; authors' calculations.

executing trades aggressively using market orders will experience larger increases in price impact than those executing trades passively using limit orders as volatility increases.

We finally examined how price impact and its sensitivity to volatility behave over time. For this purpose, we constructed a Hidden Markov Model that relates price impact to volatility in three latent states corresponding to varying liquidity conditions. Unlike in previous studies using such models to analyze liquidity, we allowed observable variables to affect the transition probabilities. We showed that times of low market depth are associated with an increased probability of low liquidity states in future. This could reflect a more severe degree of liquidity fragility associated with greater reliance on high-speed quote replenishment when depth is relatively low.

References

T. Adrian, M. J. Fleming, and E. Vogt. An Index of Treasury Market Liquidity: Evidence from 30 Years of Limit Order Book Data. *Federal Reserve Bank of New York Staff Report Number 827*, 2023.

A. Aronovich, D. Dobrev, and A. Meldrum. The Treasury Market Flash Event of February 25, 2021. *FEDS Notes, May 14, 2021*, 2021.

M. J. Barclay and J. B. Warner. Stealth Trading and Volatility: Which Trades Move Prices? *Journal of Financial Economics*, 34(3):281–305, 1993.

J. Białkowski, S. Darolles, and G. Le Fol. Improving VWAP Strategies: A Dynamic Volume Approach. *Journal of Banking and Finance*, 32(9):1709–1722, 2008.

J. A. Bikker, L. Spierdijk, and P. J. Van Der Sluis. Market Impact Costs of Institutional Equity Trades. *Journal of International Money and Finance*, 26(6): 974–1000, 2007.

Board of Governors of the Federal Reserve System. Financial Stability Report, November, 2020.

J.-P. Bouchaud, J. D. Farmer, and F. Lillo. How Markets Slowly Digest Changes in Supply and Demand. In T. Hens and K. R. Schenk-Hoppé, editors, *Handbook of Financial Markets: Dynamics and Evolution*, pages 57–160. Elsevier, 2009.

J. Brogaard, T. Hendershott, and R. Riordan. Price Discovery without Trading: Evidence from Limit Orders. *Journal of Finance*, 74(4):1621–1658, 2019.

L. Chollete, R. Næs, and J. A. Skjeltorp. *What Captures Liquidity Risk? A Comparison of Trade and Order-Based Liquidity Factors.* Norges Bank Working Paper Number 3/2007, 2007.

T. Chordia, R. Roll, and A. Subrahmanyam. Co-movements in Bid-Ask Spreads and Market Depth. *Financial Analysts Journal*, 56(5):23–27, 2000.

T. Chordia, R. Roll, and A. Subrahmanyam. Determinants of Daily Fluctuations in Liquidity and Trading Activity. *Cuadernos de Economía*, 40(121):728–751, 2003.

T. Chordia, A. Sarkar, and A. Subrahmanyam. An Empirical Analysis of Stock and Bond Market Liquidity. *The Review of Financial Studies*, 18(1):85–129, 2005a.

T. Chordia, A. Sarkar, and A. Subrahmanyam. The Joint Dynamics of Liquidity, Returns, and Volatility across Small and Large Firms. *Federal Reserve Bank of New York Staff Reports Number 207*, 2005b.

R. Cont and J.-P. Bouchaud. Herd Behavior and Aggregate Fluctuations in Financial Markets. *Macroeconomic Dynamics*, 4(2):170–196, 2000.

R. Cont, A. Kukanov, and A. Stoikov. The Price Impact of Order Book Events. *Journal of Financial Econometrics*, 12(1):47–88, 2014.

R. Cont, M. Cucuringu, and C. Zhang. Price Impact of Order Flow Imbalance: Multi-level, Cross-sectional and Forecasting. *Working paper*, 2021.

D. Dobrev and A. Meldrum. What Do Quoted Spreads Tell Us about Machine Trading at Times of Market Stress? Evidence from Treasury and FX Markets during the COVID-19-related Market Turmoil in March 2020. *FEDS Notes, September 25, 2020*, 2020.

J. D. Farmer, L. Gillemot, F. Lillo, S. Mike, and A. Sen. What Really Causes Large Price Changes? *Quantitative Finance*, 4(4):383–397, 2004.

J. D. Farmer, A. Gerig, F. Lillo, and S. Mike. Market Efficiency and the Long-memory of Supply and Demand: Is Price Impact Variable and Permanent or Fixed and Temporary? *Quantitative Finance*, 6(02):107–112, 2006.

J. D. Farmer, A. Gerig, F. Lillo, and H. Waelbroeck. How Efficiency Shapes Market Impact. *Quantitative Finance*, 13(11):1743–1758, 2013.

M. D. Flood, J. Liechty, and T. Piontek. Systemwide Commonalities in Market Liquidity. *Office of Financial Research Working Paper Number 15-11*, 2016.

J. Fraenkle, S. Z. Rachev, and C. Scherrer. Market Impact Measurement of a VWAP Trading Algorithm. *Journal of Risk Management in Financial Institutions*, 4(3): 254–274, 2011.

C. Frei and N. Westray. Optimal Execution of a VWAP Order: A Stochastic Control Approach. *Mathematical Finance*, 25(3):612–639, 2015.

N. Hautsch and R. Huang. Limit Order Flow, Market Impact and Optimal Oder Sizes: Evidence from NASDAQ TotalView-ITCH Data. In Frédéric Abergel, Jean-Philippe Bouchaud, Thierry Foucault, Charles-Albert Lehalle, and Mathieu Rosenbaum, editors, *Market Microstructure: Confronting Many Viewpoints*. John Wiley and Sons, Ltd., 2012a.

N. Hautsch and R. Huang. The Market Impact of a Limit Order. *Journal of Economic Dynamics and Control*, 36(4):501–522, 2012b.

R. A. Korajczyk and R. Sadka. Pricing the Commonality across Alternative Measures of Liquidity. *Journal of Financial Economics*, 87(1):45–72, 2008.

A. S. Kyle. Continuous Auctions and Insider Trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.

F. Lillo and J. D. Farmer. The Long Memory of the Efficient Market. *Studies in Nonlinear Dynamics and Econometrics*, 8(3), 2004.

E. Moro, J. Vicente, L. G. Moyano, A. Gerig, J. D. Farmer, G. Vaglica, F. Lillo, and R. N. Mantegna. Market Impact and Trading Profile of Large Trading Orders in Stock Markets. *Working Paper*, 2009.

R. Næs, J. A. Skjeltorp, and B. A. Ødegaard. Liquidity at the Oslo Stock Exchange. *Working paper*, 2008.

G. Nguyen, R. Engle, M. J. Fleming, and E. Ghysels. Liquidity and Volatility in the US Treasury Market. *Journal of Econometrics*, 217(2):207–229, 2020.

M. C. Pham, H. M. Anderson, H. N. Duong, and P. Lajbcygier. The Effects of Trade Size and Market Depth on Immediate Price Impact in a Limit Order Book Market. *Journal of Economic Dynamics and Control*, 120:103992, 2020.

V. Raman, M. A. Robe, and P. K. Yadav. Electronic Market Makers, Trader Anonymity and Market Fragility. *Working paper*, 2014.

Ingmar Visser and Maarten Speekenbrink. depmixs4: an r package for hidden markov models. *Journal of statistical Software*, 36:1–21, 2010.

K. Xu, M. D. Gould, and S. D. Howison. Multi-level Order-flow Imbalance in a Limit Order Book. *Market Microstructure and Liquidity*, 4:1950011, 2018.

W. Zucchini and I. L. MacDonald. *Hidden Markov Models for Time Series: An Introduction using R*. Chapman and Hall/CRC, 2009.