True Liquidity and Fundamental Prices: US Tick Size Pilot

Rohit Allena* Tarun Chordia †‡

December 16, 2021

Abstract

We develop a big-data methodology to estimate true stock prices and liquidity, explicitly considering rounding due to the minimum tick size. We apply our method to evaluate the tick size pilot (TSP), which increased the tick size for randomly chosen stocks. While the TSP increases market-maker profits it does not improve liquidity. This is consistent with theoretical models but contrasts with existing empirical studies. Rounding-adjusted true liquidity, unlike the existing liquidity measures, captures the TSP-induced trading restrictions and the decreased inventory holdings of market-makers, validating our methodology and the accuracy of our measures. It is important to account for rounding.

Keywords: Liquidity, True Prices, Rounding, Effective Spreads, Realized Spreads, US Tick Size Pilot, Machine Learning, Structural Estimation, Variational Inference, High-Frequency Data

*C.T. Bauer College of Business, University of Houston, Houston TX 30322, email:rallena@central.uh.edu

†Goizueta Business School, Emory University, Atlanta GA 30322, email:tarun.chordia@emory.edu

‡We thank Jonathan Brogaard, Kee Chung, Terry Hendershott, Ryan Riordan, Dominique Roesch, Gideon Saar, Xin Wang, Chen Yao, Mao Ye and seminar participants at the Mid-Atlantic Research Conference, MM Online Seminars - Asia Pacific, Financial Markets and Corporate Governance Conference, Deakin University, and the University of Houston for helpful comments and suggestions.

I. Introduction

Estimating stock liquidity and price efficiency measures, such as bid-ask spreads, adverse selection costs, effective spreads, realized spreads, and price discovery, is of fundamental importance to market participants, regulators and researchers. Since stocks trade at prices rounded to the grid determined by the minimum tick size, transaction prices and quoted bid-ask spreads do not represent their true fundamental values that would exist in a market with no minimum tick size. Thus, the traditional liquidity and price efficiency measures that are not adjusted for rounding and are computed using bid quotes, ask quotes, and transaction prices would be biased.[1] A formal model and methodology for estimating the rounding-adjusted liquidity and price efficiency measures in the high frequency trading (HFT) regime is conspicuously lacking in the literature.

This paper presents a structural model for true prices and spreads that explicitly accounts for the rounding specification. Our model adapts the classical Huang and Stoll (1997), Hasbrouck (1999a,b), and Ball and Chordia (2001) models to suit the current HFT environment and to account for rounding. We model the observed transaction price as the discretized (i.e., rounded) sum of i) the unobserved fundamental price that evolves as a random walk, subject to information shocks and to price discovery through the market and limit orders, and ii) the impact of trading frictions due to inventory and order processing costs. Thus, over short horizons, the observed price is a discrete version of the sum of a permanent informational component and the transient component arising due to the trading mechanism. The true spread, which equals the continuous spread that would exist in the absence of the tick, is modeled as a transform of a Gaussian autoregressive process associated with the fundamental price and other structural variables such as the time of day, time between trades, and the size and depth of the prior trade. Consistent with Hasbrouck (1999a,b), the quoted ask equals the true ask rounded up to the nearest grid point and the quoted bid equals the true bid rounded down.

Estimating the unknown true price and spread from this structural model is highly challenging

[1]Virtually all empirical studies use the mid-point of bid and ask quotes as a proxy for the unknown true price, and use the quoted bid-ask spread as a proxy for the unknown true spread. Prominent examples include Huang and Stoll (1997), Korajczyk and Sadka (2008), and Rindi and Werner (2017), among many others.

for two main reasons. First, although the resultant model takes a bivariate state-space form, the rounding specification destroys the Gaussian structure, rendering standard methods such as Kalman Filter inapplicable. To address this concern, Hasbrouck (1999a) and Ball and Chordia (2001) set up the problem in a Bayesian framework, where they use Gibbs sampling on the 15-minute and second-level data, respectively, to estimate the posterior densities of the unknown variables. However, their method is computationally infeasible on the massive millisecond-level trade and quote (TAQ) data, as it requires repeated sampling of a large number of state variables.[2] Although their method could be employed by aggregating the millisecond data to the second-level, Holden and Jacobsen (2014) demonstrate that such an aggregation yields distorted liquidity measures.

Using recent advances in machine learning, we develop a big-data methodology that tackles both, (i) the problem of non-gaussian errors and (ii) the scalability to big data. As in Ball and Chordia (2001), we set up the problem in a Bayesian state-space framework. However, instead of drawing a massive number of repeated samples, we directly approximate the posterior densities of the true prices and spreads using a procedure known as Variational Inference (VI). The main idea is to find the best density, q^*, that is statistically closest to the true posterior density. VI, thus, turns the sampling problem into an optimization problem. The optimal density q^* is obtained by iteratively solving the first order conditions, which are derived in closed-form expressions, thus, allowing for quick estimation.

We apply our method to evaluate the tick size pilot (TSP) recently conducted by the Securities and Exchange Commission (SEC). During the TSP, the SEC increased the tick size from 1 cent to 5 cents for some randomly selected stocks over the period October 2016 through September 2018. The pilot securities are divided into one control group, C, of nearly 1200 stocks and three test groups of 400 stocks each $-$ G_1, G_2 and G_3. G_1 stocks continue to trade at a one cent tick but are allowed to quote only in five cent increments; G_2 stocks are allowed to both trade and quote only in five cent increments with a few exceptions;[3] G_3 stocks are quoted and traded in five cent increments and are subject to a Trade-at-Prohibition rule, which generally prevents price matching

[2] Gibbs sampling requires drawing about 10^{13} simulations of the millisecond-level true prices and spreads.

[3] Exceptions that permit executions in one cent increments are the (1) midpoint between the national or protected best bid and the national or best protected offer, (2) retail investor orders with price improvement of at least \$0.005 per share, and (3) negotiated trades.

by a trading center that is not displaying the best price unless an exception applies. The TSP was primarily designed to test whether an increased tick size i) would enhance market-makers' profits, thus encouraging their participation, and thereby ii) improve price discovery and liquidity in the treated stocks.

The big data methodology for estimating our structural model of true spreads and fundamental prices as applied to the TSP provides a rich set of inferences that are consistent with theory. We examine the impact of TSP on market maker profits, liqudity, true spreads, components of the true spreads, the speed of incorporation of information into prices, and the proportion of price discovery through market orders, limit orders, and new information. We also demonstrate the importance of rounding by showing that the existing studies that do not account for rounding may provide biased results.

The market maker profits are estimated based on (i) the one-minute realized spreads, (ii) the difference between quoted and true spreads, and (iii) the realized profits. We discuss each of these estimated profits in turn.

Fundamentally, the 1-minute realized spread at time t equals twice the difference between the transaction price at time t and the true price one minute later for buyer-initiated trades or twice the difference between the true price at time $t + 1$ minute and the transaction price at time t for seller-initiated trades. The TSP increased the realized spreads across all the treated stocks, constrained and unconstrained,[4] and all the increases (except for those of the G_1 unconstrained stocks) are highly statistically significant.

Conrad and Wahal (2020) show that the 1-minute realized spreads might not precisely capture market-making profits for the fast traders, as they dissipate rather quickly. Thus, we develop an instantaneous measure of market-making profits, which is the difference between the quoted spreads (defined as the difference between the ask and bid quotes) and the true spreads (defined as the difference between the true ask and bid prices), denoted as mmp. While the quoted spread represents the potential revenue per (round-trip) trade earned by the market-makers, the true

[4]Unconstrained (constrained) stocks are those whose average quoted bid-ask spreads were higher (lower) than 5 cents prior to the TSP.

spread equals the sum of all costs (i.e., inventory holding, order processing and adverse selection costs) incurred by them (Demsetz (1968), Stoll (1978), and Glosten and Milgrom (1985)). Thus, *mmp* captures per trade rounding profits of market-makers. We find that the TSP significantly increased the *mmp* across all the treated stocks. In the case of the constrained stocks, *mmp* increases because of a large increase in quoted spreads of around 3.4 cents combined with a small decrease in the true spreads. This increase in quoted spreads occurred in spite of the decline in the adverse selection and the inventory cost components of the true spread suggesting that the increase in the quoted spreads is driven entirely by the 5 cent tick size constraint imposed by the TSP. In the case of the unconstrained stocks, *mmp* increases mainly due to a large decrease in the true spreads ranging from 3 to 3.6 cents (along with a small increase in the quoted spreads for the G_3 stocks), which is in turn driven by a large decrease in the inventory cost (along with a far smaller decrease in the adverse selection cost) faced by the liquidity providers.

The decreases in the adverse selection and the inventory cost components of the true spreads are consistent with the model of Goettler, Parlour, and Rajan (2005, 2009). In the case of constrained stocks, the true spreads are much lower than the quoted spreads due to the large binding tick size. The rounding driven artificially large quoted spreads reduce the adverse selection and inventory risk faced by market makers as the fundamental price is more likely to remain within the quoted spreads. This lowers the adverse selection and inventory risk components of the true spread. To extract the large rounding profits, the high frequency traders (HFTs), given their speed advantage, provide liquidity by submitting fast limit orders to position themselves in the front of the limit order queues. These fast orders crowd out the relatively slow informed traders, who incorporate new information into stock prices. Although the informed traders could potentially undercut limit orders of HFTs, it is expensive due to the large tick size and often not possible as the minimum tick is binding. Thus, queue competition rather than price competition prevails in the tick-constrained environment resulting in larger depths at the quotes (O'Hara, Saar, and Zhong (2019)), which is what we find. This queueing equilibirum is consistent with the model of Li, Wang, and Ye (2020) as well as the results in Yao and Ye (2018).

In the case of the unconstrained stocks, undercutting of the resting limit orders is possible. As a result, due to the increased tick size informed traders strategically cut resting limit orders when it is profitable for them to do so. Since it is the informed traders, and not the market makers, who are likely to provide liquidity at the inside quotes, they are less likely to be subject to inventory risk as they do not carry inventory. The uninformed market makers have a lower incentive to provide the inside quotes because of the strategic undercutting by the informed traders, which would adversely select the uninformed market makers. Further, given their information advantage, the informed traders are also less subject to adverse selection. This is the undercutting equilibrium proposed by Goettler et al. (2005, 2009) and it predicts lower adverse selection and inventory risk costs for the informed liquidity providers, which is what we find. We also find lower depths as in O'Hara et al. (2019).

The *mmp* captures the rounding profits earned by market makers from a round-trip trade. However, for a given trade, market makers earn rounding profits from either the ask or the bid quote depending on whether the trade is a buy or a sell. The realized profit measure is defined as transaction price minus the true ask if the trade is a customer buy (market-maker sell) and the true bid minus the transaction price if the trade is a customer sell (market-maker buy). The realized profits follow the same pattern as that of the *mmp*. In general, the realized profits are less than half those of the *mmp*. This suggests that the rounding driven supply of liquidity is not as profitable as expected by naively comparing quoted and true spreads because the liquidity demanders are able to trade at prices where the impact of rounding is the lowest. Thus, some market participants are able to ascertain the true prices and spreads. Since traders use sophisticated algorithms to understand true prices and liquidity, we, as econometricians, should also use sophisticated big-data methods to back out these true measures. This further motivates our use of sophisticated big-data methods to estimate true prices and spreads.

Next, we turn to liquidity and the transaction costs faced by investors as proxied by the effective spreads. As we noted above, the TSP increased the quoted spreads, especially for the constrained stocks. The effective spread, which quantifies the cost of trading the stock, is fundamentally defined as twice the absolute difference between the transaction price and the true fundamental price of the stock (Bessembinder (2003)). The TSP mechanically increased the effective spreads across all the treated stocks and all the increases (except for those of the G_1 unconstrained stocks) are highly statistically significant. Not surprisingly the increase in the effective spreads is larger for the constrained stocks due to the binding minimum tick.

The SEC's goals in implementing the TSP of testing whether it is successful in increasing market maker profits and liquidity has mixed results. The TSP does increase market maker profits but it lowers liquidity and increases transactions costs for investors. The speed of incorporation of information into prices also decreases for the constrained stocks but it increases for the unconstrained stocks (albeit insignificantly so) possibly due to the undercutting by the informed traders. Moreover, the TSP reduces the proportion of price discovery through new information (relative to price discovery through limit and market orders), pointing to a decline in price efficiency, at least for the constrained stocks. It is not surprising to find a decline in price efficiency in the constrained stocks as the informed traders are crowded out in the queueing equilibrium. Due to the undercutting equilibrium in the unconstrained stocks, informed traders have more of an incentive to obtain fundamental information, which they incorporate into prices through limit orders, thereby decreasing the proportion of price discovery through new information.

Existing empirical studies on the TSP, including Rindi and Werner (2017), Comerton-Forde, Grégoire, and Zhong (2019), Chung, Lee, and Roesch (2019), and Albuquerque, Song, and Yao (2020), use various liquidity measures that do not account for rounding. In addition, they make inferences based on mid-quotes and quoted spreads as proxies for true prices and spreads, respectively. We show that these inferences are biased. For instance, (i) In contrast to our finding that effective spreads increased significantly for all treated stocks, except for the G_1 unconstrained stocks, Rindi and Werner (2017), Chung et al. (2019), and Albuquerque et al. (2020) find no increase in the effective spreads for the unconstrained stocks. (ii) In contrast to our finding that realized spreads increased significantly for all treated stocks, except for the G_1 unconstrained stocks, Rindi and Werner (2017) and Chung et al. (2019) find no increase in the realized spreads for the unconstrained stocks. (iii) We find that the adverse selection and the inventory cost components of the true spreads decrease due to TSP but they increase significantly for the constrained stocks when rounding is not accounted for as in the model of Huang and Stoll (1997).

Are this paper's liquidity measures and inferences more valid than the existing empirical studies that do not account for rounding? We show that our inferences indeed are more accurate based on two key empirical observations. First, our liquidity measures capture the differences in the quoting and trading restrictions imposed by the TSP across groups G_1, G_2, G_3. Recall that stocks of G_1 and G_2 are allowed to trade in 1 cent increments in all cases and some cases, respectively. Thus, it is expected that the increases in the market-maker profits and transaction costs per trade will rank in the order of $G_1 < G_2 < G_3$. Our liquidity measures do indeed conform to this order, whereas the measures from the existing studies do not. Second, the SEC made publicly available the aggregate daily inventory holdings of all the designated market-makers around the TSP period. This data suggests that the TSP significantly decreased the inventory costs of the market-makers. Our inventory components of the spreads precisely capture this TSP-driven decreased inventory costs, whereas the existing inventory component measures (Huang and Stoll (1997)) do not.

Overall, we develop a big-data method to estimate rounding-adjusted true stock prices and liquidity. We find that the TSP increased market-maker profits but decreased liquidity. This result contrasts existing empirical studies, but is consistent with the theoretical studies that do account for rounding. Our liquidity measures, unlike the existing measues, capture key TSP-induced trading restrictions and increased market-making profits, suggesting that our inferences are more accurate.

Related Literature. Methodologically, our paper relates to Hasbrouck (1999a) and Ball and Chordia (2001), who provide Bayesian methods to estimate true liquidity after accounting for rounding. We make two contributions relative to their work. First, their procedures are computationally infeasible on the millisecond-level data. For example, Hasbrouck (1999a) extracts true liquidity for only one specific stock at the 15-minute level and Ball and Chordia (2001) extract true prices and spreads for seven large stocks at the one second frequency. In contrast, we estimate true prices and spreads for a large cross-section (nearly 2400) of stocks over four months of the millisecond-level TAQ data. Second, our model allows for price discovery through limit orders, whereas the other two do not. Our model specification is consistent with the recent empirical evidence by Brogaard, Hendershott, and Riordan (2019), who attribute a large fraction of price discovery to limit orders.

Our paper also relates to Hagstromer (2020), who proposed computing effective spreads with two different proxies for the true prices. One with the depth-weighted bid-ask quotes and the other with the micro-price of Stoikov (2018). Our paper differs from Hagstromer (2020) in several dimensions. First, Hagstromer's method only yields estimates of true prices and effective spreads, whereas our procedure can estimate a variety of other important liquidity measures, such as true spreads, spread components, and the proportions of price discovery. Second, the validity of Hagstromer's depth-weighted effective spreads relies on the key assumption that informed traders submit market orders. However, we show in this paper that the majority of price discovery happens through limit orders. As a result, we find that, unlike this paper's effective spreads, the depth-weighted effective spreads do not capture the differences in trading restrictions imposed by the TSP, suggesting that our effective spread measures are more accurate.

This paper also relates to the growing literature on the applications of machine learning methods in finance. Whereas studies, such as Gu, Kelly, and Xiu (2020) and Chinco, Clark-Joseph, and Ye (2019) apply machine learning to empirically identify the best models or predictors, we use it to estimate a well-defined structural model. Thus, our inferences are economically interpretable, which is usually difficult with machine learning. Our VI methodology is general and could be applied to tackle other questions involving big-data in finance. For example, Allena (2021) exploits VI to estimate standard errors of expected return predictions from neural networks.

II. Model

Our model is a generalization of the Ball and Chordia (2001) and Huang and Stoll (1997) models and is designed to accomodate the current high frequency trading (HFT) environment. The observed transaction price P_t is modeled as

$$P_t \equiv [p_t^{NR}]_{Round} = [m_t + (1-\lambda)s_t Q_t/2]_{Round}, \tag{1}$$

where p_t^{NR} is the nonrounded price at time t; m_t is the fundamental price of the security at time t, immediately after a trade; Q_t is a trade indicator for buyer/seller classification of trades and is $+1$ if the trade is buyer initiated, -1 if the trade is seller initiated, and 0 if we are unable to sign the trade; λ is the adverse selection component of the spread; and s_t is the true spread that would obtain in a market with continuous prices i.e., a zero tick size. The notation $[.]_{Round}$ indicates rounding onto the tick grid. Thus, the observed transaction price is a result of rounding or discretization of the sum of the fundamental price and the inventory and order processing component of the spread. Note that in the presence of rounding, the disturbances in observed price changes are not Gaussian. Most market microstructure models ignore rounding, and, thus are unlikely to be correctly specified, especially if the rounding is severe.

The fundamental price (m_t) updates the past price (m_{t-1}) by incorporating any new information contained in the market orders, limit orders and other sources of publicly available information. We assume that m_t evolves as follows:

$$
\begin{aligned}
m_t = m_{t-1} + & \overbrace{\lambda \frac{s_t Q_t}{2}}^{\text{price discovery through market orders}} + \overbrace{\epsilon_t}^{\text{price discovery through public info}} + \\
& \underbrace{\lambda_1 (\Delta A_t)\, \mathrm{D}_t^A + \lambda_2 (\Delta B_t)\, \mathrm{D}_t^B + \lambda_3\, (\Delta D_t^A)\, I_{\Delta A = 0} + \lambda_4 (\Delta D_t^B)\, I_{\Delta B = 0}}_{\text{price discovery through limit orders}},
\end{aligned} \tag{2}
$$

where $\{\epsilon_t, t = 1, 2, \ldots, T\}$ are *i.i.d* $N(0, \sigma_\epsilon^2)$ and represent information shocks. The second term in equation (2) is the half fraction of spread attributable to adverse selection and represents price discovery through the market orders. The final term is the contribution of limit orders to the price discovery, where A_t (B_t) are the NBBO ask (bid) quotes and D_t^A (D_t^B) are the corresponding depths at time t. $\Delta(X_t)$ denotes the first order difference $X_t - X_{t-1}$ and I is an indicator variable for when $\Delta A = 0$ or $\Delta B = 0$.

A key distinction of our model is that we allow for price discovery through limit orders. Ball and Chordia (2001) and Huang and Stoll (1997) allow for price discovery only through the adverse selection component of the market orders and through information shocks. Our specification is consistent with the recent empirical evidence of Brogaard et al. (2019) who, due to the presence of HFTs, attribute a majority of the price discovery to limit orders. We use publicly available information including the best ask and bid quotes, corresponding depths, and their first order differences to capture price discovery through limit orders.

Given a tick-size regime, we model the dynamics of true spreads as a first order logarithmic auto-regression with additional structural variables as in Ball and Chordia (2001) :

$$ln(s_t) = \alpha + \beta ln(s_{t-1}) + \delta ln\frac{V_{t-1}}{D_{t-1}} + \tau Time_{t-1} + d_1 BEG_t + d_2 END_t + e_t, \tag{3}$$

where $\{e_t, t = 1, 2, \ldots, T\}$ are $i.i.d$ $N(0, \sigma_e^2)$, V_{t-1} is the volume of stock transacted at the previous trade, D_{t-1} is the corresponding bid or ask depth, $Time_{t-1}$ is the time, in seconds between the last trade and the one before it and BEG_t (END_t) is an indicator variable denoting the first (last) hour of the trading day.

The regression specification is consistent with the empirical evidence in Chordia, Roll, and Subrahmanyam (2001) that shows how the relative size of trade to depth on the previous transaction possibly impacts the ensuing spread at the current transaction. The dummy variables capture the intraday seasonalities and the use of lagged time between trades is motivated by Easley and O'Hara (1992), who suggest that absence of trades may provide information about the occurence of information events. Note that, due to rounding, the quoted spread cannot be modeled as an auto-regressive process with Gaussian errors. However, the (log) true spread lies on the real line and is modeled as in equation (3).

We denote $x_t = m_t - \lambda s_t Q_t/2$ and $\gamma_t = log(s_t)$ for algebraic convenience. Combining the above equations we have the following econometric model:

$$P_t \equiv [p_t^{NR}]_{Round} = [x_t + s_t Q_t/2]_{Round}, \tag{4}$$

$$x_t = x_{t-1} + \lambda \frac{s_{t-1}Q_{t-1}}{2} + +l_1 L_{1t} + l_2 L_{2t} + l_3 L_{3t} + l_4 L_{4t} + \epsilon_t, \tag{5}$$

$$\gamma_t = \alpha + \beta\gamma_{t-1} + d_1 D_{1t} + d_2 D_{2t} + d_3 D_{3t} + d_4 D_{4t} + e_t, \tag{6}$$

where the independent variables are, $L_{1t} = (\Delta A_t)\,\mathrm{D}_t^A$, $L_{2t} = (\Delta B_t)\,\mathrm{D}_t^B$, $L_{3t} = (\Delta D^A)$ $I_{\Delta A_t=0}\,Q_t$, $L_{4t} = (\Delta D^B)\,I_{\Delta B_t=0}\,Q_t$, $D_{1t} = ln\frac{V_{t-1}}{D_{t-1}}$, $D_{2t} = Time_{t-1}$, $D_{3t} = BEG_t$ and $D_{4t} = END_t$. Denoting $z_t = \{x_t, \gamma_t\}$, the system is expressed as the following first order vector autoregression (VAR(1)) model:[5]

$$z_t = \mu_t + A_t z_{t-1} + \epsilon_t. \tag{7}$$

The true spreads and fundamental prices (z_t) are not observable, whereas the transaction prices and quoted spreads that are discretized to the nearest grid are available. We adhere to the following discretization process for the observed transaction prices and quoted spreads:

1. If the observed price, P_t, is at the ask (bid) then we assume that the nonrounded price, p_t^{NR}, has been rounded up (down) to the nearest tick. Furthermore, the bid (ask) price is assumed to have been rounded down (up). Thus, for a trade at the ask, $x_t + s_t/2 \in [P_t - tick, P_t]$ and $x_t - s_t/2 \in [B_t, B_t + tick]$. Similarly, for a trade at the bid, $x_t - s_t/2 \in [P_t, P_t + tick]$ and $x_t + s_t/2 \in [A_t - tick, A_t]$.

2. If the trade is a customer buy, $Q_t = +1$, and the price is not the same as the ask, $P_t \neq A_t$, then $x_t + s_t/2 \in [P_t - tick, P_t]$ and $x_t - s_t/2 \in [B_t, B_t + tick]$.

3. If the trade is a customer sell, $Q_t = -1$, and the price is not the same as the bid, $P_t \neq B_t$, then $x_t - s_t/2 \in [P_t, P_t + tick]$ and $x_t + s_t/2 \in [A_t - tick, A_t]$.

Thus, at each time point t, we have the following information

$$\begin{aligned} x_t + s_t/2 &\in I_{1t}, \\ x_t - s_t/2 &\in I_{2t}, \end{aligned} \tag{8}$$

where, I_{1t} and I_{2t} indicate the intervals of length equal to the tick size, that each linear

[5] Note that our VAR model uses only lagged information.

functional of the state variable must lie within. In other words, the observed information places the adjusted fundamental price plus the half-spread in one interval of length equal to the tick size and places the adjusted fundamental price minus the half-spread in another length equal to the tick size. We use the above specification with a uniform *tick* of 1 *cent* across all the treated stocks in groups G_1, G_2, G_3 and the control group C during the non-pilot regime.

For the periods considered during the TSP, we adapt the following tick rule:

1. Given that the SEC restricts stocks in group G_1 to quote only in 5 *cents* but are allowed to trade in 1 *cent*, we use a *tick* of 5 *cents* for rounding the quoted bid-ask spreads, and a *tick* of 1 cent for the transaction prices. This is also equivalent to rounding all the trades at the ask or the bid to 5 *cents*, and rounding the remaining trades with transaction prices between the bid-ask quotes to 1 *cent*.

2. We use a *tick* of 5 *cents* for stocks in the groups G_2, G_3 since these are restricted to quote and trade only in 5 *cents*. However, for G_2 stocks, when we observe transaction prices that are not on the five cent tick grid, we round to one cent.

3. Lastly, we use a *tick* of 1 *cent* for stocks in the control group (C) because they continue to both trade and quote in 1 *cent* increments.

Summarizing the set of observed values at time period t as $Y_t = \{P_t, A_t, B_t, tick\}$, our interest lies in estimating the hidden state variables (z_t) that includes true spreads and fundamental prices, and the set of parameters $\Theta = \{\lambda, \beta, l_1, l_2, l_3, l_4, d_1, d_2, d_3, d_4, \sigma_\epsilon^2, \sigma_\eta^2\}$. We cast our econometric model in equations (7) and (8) into the state space framework with equation (7) as the transition equation and (8) as the measurement equation. The rounding mechanism embedded in the measurement equation destroys the Gaussian structure and the time series independence of the errors, rendering standard estimation methods (Kalman Filtering, Kalman (1960)) invalid.

Thus, recognizing the challenges posed by the massive millisecond-level TAQ data and the rounding specification, we develop a computationally scalable VI method to estimate true prices and spreads in the following section.

III. Methodology: Variational Inference (VI)

The goal is to estimate the state variables $z_t = \{x_t, \gamma_t\}$ and the parameters Θ, having observed $\{Y_t\}$ and other explanatory variables such as $\{Q_t\}$ and $\{L_{it}, D_{it}\}_{i=1}^4$. We set up the estimation problem in a Bayesian framework due to the discreteness and rounding structure of the observed variables. The premise of a Bayesian framework is to place priors on the latent variables and parameters, and estimate the posterior density of latent variables and parameters given the priors and the observed data,

$$P(z_{t=1}^T, \Theta | Y_{t=1}^T) \propto P(Y_{t=1}^T | z_{t=1}^T, \Theta) P(z_{t=1}^T | \Theta) P(\Theta), \tag{9}$$

where $P(\Theta)$ is the prior density of the parameters, $P(Y_{t=1}^T | z_{t=1}^T, \theta)$ is the conditional likelihood of $Y_{t=1}^T$ given $z_{t=1}^T$ and Θ, and $P(z_{t=1}^T | \Theta)$ is the conditional likelihood of $z_{t=1}^T$ given Θ.

The calculation of joint posterior is generally intractable. Considering the Markovian structure of latent variables and the rounding specification of our model, the conditional posterior densities of each variable or parameter given the other parameters, variables and the data could be easily expressed as known density functions. Building on this insight, Ball and Chordia (2001) conduct Gibbs sampling by drawing a large number of samples from the conditional densities $P(\Theta | z_{t=1}^T, Y_{t=1}^T)$ and $P(z_t | z_{\sim t}, \Theta, Y_{t=1}^T)$, for $t = 1, 2, \ldots T$, where $z_{\sim t}$ is the set of all latent variables excluding z_t. However, this procedure is computationally infeasible on the massive millecond-level TAQ data. For example, Ball and Chordia (2001) extract fundamental prices and true spreads for only seven large and mid-cap stocks using the second-level transactions data. This data is further restricted to a maximum of 14,000 stock-level transactions over 1 month sample period. In contrast, an average stock with four months of *millisecond-level* TAQ data has over 6×10^5 transactions. Thus, evaluating the TSP across 2400 stocks via Gibbs sampling requires drawing simulations of about 29×10^8 ($\sim 2 \times 2400 \times 6 \times 10^5$) latent variables. Furthermore, these simulations would have to be repeated a large number of times (10,000 times in Ball and Chordia (2001)) until all the latent variables and parameters converge.

Our VI method directly approximates the posterior density by solving an optimization problem and bypasses the challenge of drawing large number of repeated samples. In particular, to approximate the posterior density $P(z_{t=1}^T, \Theta | Y_{t=1}^T)$, we consider a family of known densities $\mathcal{Q}$ over the latent variables $(z_{t=1}^T)$ and the parameters Θ. Each density q $(\in \mathcal{Q})$ in the family is a candidate approximation for the true posterior. The premise of our methodology is to find the best density in the family, q^* $(\in \mathcal{Q})$, that is (statistically) closest to the true posterior density in terms of the Kullback-Leibler (KL) divergence,

$$
\begin{aligned}
q^*(\{z_t\}_{t=1}^T, \Theta) &= \arg \min_{q(\{z_t\}_{t=1}^T, \Theta) \in \mathcal{Q}} KL\left[q(\{z_t\}_{t=1}^T, \Theta) || P\left(\{z_t\}_{t=1}^T, \Theta | Y_{t=1}^T\right)\right] \\
&= \arg \min_{q(\{z_t\}_{t=1}^T, \Theta) \in \mathcal{Q}} E\left[\log\left(q(\{z_t\}_{t=1}^T, \Theta)\right)\right] - E\left[\log\left(P\left(\{z_t\}_{t=1}^T, \Theta | Y_{t=1}^T\right)\right)\right], \quad (10)
\end{aligned}
$$

where Kullback-Leibler (KL) distance between two densities quantifies how much the second density is different from the first, and all expectations are taken with respect to the considered density over the latent variables and parameters, $q(.)$. Finally, we approximate the posterior with the optimized member of the family $q^*(.)$. Variational inference thus turns the sampling problem into an optimization problem. The key is to consider a generous family of densities $\mathcal{Q}$ such that a member of the family closely approximates the true posterior, but is simple enough for solving the optimization problem.

Minimizing the KL objective appears to be not possible since the true posterior density, $P(\{z_{t=1}^T, \Theta | Y_{t=1}^T\})$ is not known. However, a useful decomposition of the second term in equation (10) shows that this minimization objective is solvable despite the absence of true posterior density. The decomposition is as below:

$$
\log\left(P(z_{t=1}^T, \Theta | Y_{t=1}^T)\right) = \log\left(P(z_{t=1}^T, \Theta, Y_{t=1}^T)\right) - \log\left(P(Y_{t=1}^T)\right), \quad (11)
$$

where the first term of equation (11) is the joint density of the latent variables, parameters and the observed data that can be computed using the priors on the parameters and the latent variables; and the likelihood of the data given the parameters and the latent variables. The second term is the marginal likelihood of the observed data that involves integrating the likelihood function with respect to the priors on parameters and the latent variables. Although not computable, this term is free of parameters and the latent variables, and thus is a constant with respect to any density $q(.)$, over the latent variables and parameters. Therefore, the minimization objective involves only the known priors and likelihood functions, which are solvable. In particular, it is equivalent to maximizing the popularly known objective, Evidence Lower Bound ($ELBO$), which is defined below:

$$ELBO(q) = E\left[\log\left(P\left(\{z_t\}_{t=1}^T, \Theta, Y_{t=1}^T\right)\right)\right] - E\left[\log\left(q(\{z_t\}_{t=1}^T, \Theta)\right)\right]. \tag{12}$$

Blei, Kucukelbir, and McAuliffe (2017) outline a procedure that addresses two key questions for a set of specialized models that belong to the exponential family: **i**) what family of densities to consider for the approximation of the latent variables? **ii**) how to obtain the optimal density in the family that best approximates the true posterior? We generalize their theory and derive a procedure for approximating the posterior in the context of discreteness using two main results:

1. The likelihood function of observed prices and spreads given true spreads, prices and parameters is a truncated bivariate normal density.

2. Truncated bivariate normal densities belong to exponential set of family.

In what follows, we lay out the procedure for choosing a family of densities, $\mathcal{Q}$ to approximate the joint posterior of the latent variables and parameters, and obtaining the optimal density $q^* \in \mathcal{Q}$ that best approximates the true posterior.

A. Family of Densities for Approximation

Our idea is to approximate the posterior density of latent variables given observed variables by solving an optimization problem. We use a family of densities over the latent variables, parametrized by "variational parameters". The optimization finds the member of this family, that is, the setting of "variational parameters", which is closest to the true posterior density of latent variables.

We consider a specific family of densities, where the latent variables and parameters are independent. These are popularly known as mean-field densities and each of its candidate density is of the form:

$$q(z_{t=1}^T, \Theta|\Phi) = \Pi_{t=1}^T q(z_t; \Phi_{z_t}) \Pi_{i=1}^4 q(l_i; \Phi_{l_i}) q(d_i; \Phi_{d_i}) q(\lambda; \Phi_\lambda) q(\beta; \Phi_\beta) q(\alpha; \Phi_\alpha), \tag{13}$$

where $\Phi = \{\{\Phi_{z_t},\}_{t=1}^T, \{\Phi_{l_i}, \Phi_{d_i}\}_{i=1}^4, \Phi_\lambda, \Phi_\beta, \Phi_\alpha\}$ are the "variational parameters" governing approximate densities $q(.|\Phi)$. Our goal is then to find the optimal density $q^*(.|\Phi^*)$, or the variational parameters Φ^*, such that the density $q^*(.|\Phi^*)$ is closest to the true posterior $P(z_{t=1}^T, \Theta|y_{t=1}^T)$. Equivalently, q^* is the solution to the below optimization problem:

$$\begin{aligned} q^*(z_{t=1}^T, \Theta|\Phi^*) &= \Pi_{t=1}^T q^*(z_t; \Phi_{z_t}^*) \Pi_{i=1}^4 q^*(l_i; \Phi_{l_i}^*) q^*(d_i; \Phi_{d_i}^*) q^*(\lambda; \Phi_\lambda^*) q^*(\beta; \Phi_\beta^*) q^*(\alpha; \Phi_\alpha^*) \\ &= \arg\min KL(q(z_{t=1}^T, \Theta)||P(z_{t=1}^T, \Theta|y_{t=1}^T)) \end{aligned}$$

Before describing the procedure for obtaining the optimal density q^*, it is worth highlighting few properties of the mean-field family of densities. Note that the candidate densities $q \in \mathcal{Q}$, assumes that all the parameters and hidden-states are time-independent. However, the state variables and the parameters in the true posterior are not time-independent. Given that we are minimizing the Kullback-Leibler distance, this simple approximation, however works well in approximating the true marginal posteriors, and thus means and variances of the individual state variables. For example, Wang and Blei (2018) prove that, asymptotically, VI with mean-field families is a theoretically sound approximate inference procedure for the marginal densities, even though it tends to underestimate the covariances of different variables' joint posterir densities. Because we estimate fundamental prices, true spreads and parameters (Θ) using means and standard deviations of marginal posterior densities (rather than joint densities) of respective variables and parameters, mean-field approximations aptly serves our purpose. We also demonstrate the success of the mean-field approximations in our context using Monte-Carlo simulations in the next section.

B. Estimating the Optimal Density Function

Estimating the optimal mean-field density, $q^*(.|\Phi^*)$ that is closest to the true posterior is not straight forward and does not have a closed form solution. However, Blei et al. (2017) derive a useful result, which shows how the optimal density of an individual latent variable given the optimal densities of all other latent variables could be easily obtained. Therefore, $q^*(.|\Phi^*)$ can be estimated by starting with some initial guesses on the optimal densities and recursively updating the optimal density of each variable given the optimal densities of other variables. We first state the fundamental result on mean-field approximations due to Blei et al. (2017):

Proposition 1. *Given observations $\{Y_{t=1}^T\}$, the optimal mean-field density of a hidden variable $z_t, q^*\left(z_t|\Phi_{z_t}^*\right)$ given the optimal densities of other hidden variables $z_{\sim t}, q^*\left(z_{\sim t}|\Phi_{z_{\sim t}}^*\right)$ and the parameters Θ, $q^*\left(\Theta|\Phi^*\right)$ is proportional to*

$$q^*\left(z_t|\Phi_{z_t}^*\right) \propto \exp\left(E_{q^*\left(z_{\sim t}|\Phi_{z_{\sim t}}^*\right)q^*(\Theta|\Phi^*)}\log\left[f\left(Y_{t=1}^T|z_{t=1}^T,\Theta\right)f(z_{t=1}^T|\Theta)f(\Theta)\right]\right), \tag{14}$$

where $f\left(Y_{t=1}^T|z_{t=1}^T,\Theta\right)$ is the conditional density of the observed variables $Y_{t=1}^T$ given the hidden state variables $z_{t=1}^T$ and the parameters Θ; $f(z_{t=1}^T|\Theta)$ is the conditional density of hidden state-variables given the parameters Θ; and $f(\Theta)$ is the prior density of the parameters. Note that the expectation in the above equation is taken with respect to the optimal densities of excluded hidden variables $q^*(z_{\sim t}|\Phi_{z_{\sim t}}^*)$ and the parameters $q^*(\Theta|\Phi^*)$. Similarly, the optimal density of parameters Θ, $q^*(\Theta|\Phi^*)$ is given by

$$q^*\left(\Theta|\Phi^*\right) \propto \exp\left(E_{q^*\left(z_{t=1}^T;\Phi_{z_{t=1}^T}^*\right)}\log\left[f\left(Y_{t=1}^T|z_{t=1}^T,\Theta\right)f(z_{t=1}^T|\Theta)f(\Theta)\right]\right). \tag{15}$$

Here, the expectation is taken with respect to the optimal densities of state variables, $q^*\left(z_{t=1}^T;\Phi_{z_{t=1}^T}^*\right)$.

Starting with some initial values for the "variational parameters" $\{\{\Phi_{z_t}^*\}_{t=1}^T,\Phi^*\}$, we can recursively update the "variational parameters" or the optimal densities given the other variational parameters using (14) and (15) until convergence. Before deriving the update equations, it is worth pointing out the similarities of this methodology with the Gibbs Sampling approach of Ball and Chordia (2001). While a large number of samples are recursively drawn in Gibbs Sampling from the conditional posteriors of a variable given other variables and parameters, Variational Inference directly updates the moments of marginal posteriors, and thus bypasses the challenge of drawing a large number of samples.

C. Derivations of Updates

We derive the equations for updating the optimal density of a parameter or a hidden variable given other variables and parameters in propositions 2-6 in the appendix. Then we estimate the posteriors of the true prices, spreads and other parameters, including the adverse selection component (λ), by recursively updating the variational densities given in these propositions (equations (27), (29), (31), (32), (33)) until all the densities converge. Overall, the algorithm for approximating the true posterior is given below :

1. Set initial values of variational parameters Φ, Φ_{z_t} for approximating densities of parameters, Θ and state-variables, $z_{t=1}^{T}$ respectively.
2. Update variational density of parameters ($q^*(\Theta)$) using equations (27), (29), (31), (32).
3. Update variational density of state variables ($q^*(z_t)$) using equation (33).
4. Compute Evidence Lower Bound ($ELBO$) using equation (12), and steps 2 and 3.
5. Repeat steps 2 to 4 until $ELBO$ convergence.
6. Approximate the true posterior with $q^*(z_{t=1}^{T}, \Theta)$.

IV. Performance of Our Methodology: Simulation Evidence

Using Monte-Carlo simulations, we show that our methodology outperforms existing methods in estimating true prices and liquidity in terms of both precision and latency. For example, simulations suugest that our method reduces the mean squared errors of true prices by 1.7 times and true spreads by 11 times, relative to the existing proxies.

A. Accuracy of the methodology

We first evaluate the accuracy of our methodology by examining whether the estimated parameters and variables that include true spreads, effective spreads and market-makers profits are close to the respective true simulated values. We then compare our methodology with the existing procedures that naively estimate these variables with the observed transaction prices and quoted spreads, without adjusting for rounding.

Given a set of regression coefficients $\{\lambda, \alpha, \beta, l_1, l_2, l_3, l_4, d_1, d_2, d_3, d_4\}$ and other explanatory variables $\{Q_t, L_{1t}, L_{2t}, L_{3t}, L_{4t}, D_{1t}, D_{2t}, D_{3t}, D_{4t}\}$, we simulate fundamental prices (x_t) and true spreads (s_t) using the VAR(1) specification in equations (5) and (6). The regression coefficients $\{\lambda, \alpha, \beta, \ldots, d_4\}$, and the explanatory variables $\{Q_t, L_{1t}, \ldots, D_{4t}\}$ are calibrated to match their empirical counterparts for a given stock. For example, to simulate fundamental prices and true spreads for the stock of Florida Community Bank (FCB), we sign the trades during the sample period as buys and sells to obtain Q_t; we use the product of change in its best ask quote and its depth at the ask ($\Delta A_t \times D_t^A$) as L_{1t}, and similarly for other variables $\{L_{2t}, L_{3t}, \ldots, D_{4t}\}$. For the parameters $\{\lambda, \alpha, \beta, \ldots, d_4\}$, we use respective sample estimates (OLS) of coefficients in the VAR(1) specification of (5) and (6), with mid-quotes as x_t, and half spreads ($(Ask_t - Bid_t)/2$) as s_t.

After simulating the fundamental prices and true spreads, we use the rounding rule (8) to obtain the observed ask quotes of $A_t = [x_t + s_t/2]^{up}$, and the bid quotes of $B_t = [x_t - s_t/2]^{down}$. Without loss of generality, for a grid size of 5 cents, we specify that the rounded prices and quotes are exact multiples of 5 cents. For example, if the true simulated ask quote of a stock ($x_t + s_t/2$) is \$32.3245 cents, the observed ask quote (A_t) is rounded up to 32.35. Similarly, if the true bid quote ($x_t - s_t/2$) is \$31.3212, then B_t is rounded down to \$31.30. Using these simulated rounded values of $\{A_t, B_t, P_t\}$, and the explanatory variables $\{q_t, L_{1t}, L_{2t}, L_{3t}, L_{4t}, D_{1t}, D_{2t}, D_{3t}, D_{4t}\}$, we implement our methodology to estimate the latent variables $\{x_t, \gamma_t\}$ and the parameters $\{\lambda, \alpha, \beta, l_1, l_2, l_3, l_4, d_1, d_2, d_3, d_4\}$, and check whether the methodology recovers the true simulated variables and parameters.

Table 1 showcases the performance of our methodology, where we conduct independent Monte-Carlo simulations calibrated to three randomly selected stocks, Florida Community Bank (FCB), Boston Beer Company (SAM) and AMC Entertainment Holdings (AMC). The first column represents the true simulated values, while the second and third are the estimated values using our methodology and existing procedures, respectively. We find that our methodology performs well by noting that the difference between average simulated fundamental prices and estimated fundamental prices is 0.07, 0.11 and 0.12 cents and the difference between the average simulated true effective spreads and estimated true effective spreads is 0.02, 0.05, and 0.01 cents for FCB, SAM, and AMC, respectively. We further establish the success of our methodology by showing that other variables such as price impact (λ), market maker profits (mmp) and the posterior variance estimates also closely match the corresponding true simulated values.

Simulations also show that the existing liquidity measures without the rounding adjustment are highly biased. For example, when the fundamental prices and true spreads are simulated under parameters calibrated to match those of SAM, the effective spread is biased by 20%. Similarly, when the parameters are calibrated to match FCB (AMC), the squared sum of error of naively estimating the true spreads with the quoted spreads is 11-times (16-times) more than the squared sum of error of our estimated true spreads. Thus, the simulation evidence not only validates our methodology but also underlines the biases in existing procedures in estimating fundamental microstructure variables.

B. Scalability of the methodology: ELBO Convergence

We also assess the scalability of our methodology by examining the number of iterations required for the algorithm to converge. Figure 1 shows that for AMC, with nearly 4×10^5 transactions, the algorithm requires only 6 iterations (epochs - in machine learning parlance) and less than 2 minutes to converge on a person computer with an I7 - (790 CPU, 3.6GHz) processor and 16 GB RAM.

V. Empirical Results

The Securities and Exchange Commission (SEC) conducted the TSP over the period October 2016 through September 2018, increasing the tick size from 1 cent to 5 cents for a randomly selected sample of stocks.[6] The TSP was primarily designed to test whether an increased tick size i) would enhance market-makers' profits (mmp), thus encouraging their participation, and thereby ii) improve price discovery and liquidity in the treated stocks. In this section, we use our true liquidity and price measures to provide empirical evidence about the two SEC goals, of increasing market maker profits and improving price discovery and liquidity. We also discuss how and why the existing liquidity measures that do not account for rounding deliver contrasting inferences.

In addition, this section validates that our rounding-adjusted inferences are more accurate than the existing studies by showing that our true liquidity estimates, unlike the existing measures, capture the TSP-driven trading restrictions and the decreased market-makers' inventory costs. We also document several patterns in market-maker profits and depths that indicate how sophisticated traders, such as HFTs, are able to trace out these true stock prices and spreads, further validating our rounding-adjusted liquidity measures.

A. Data

Our sample consists of daily millisecond-level TAQ across all stocks included in the TSP. The data spans over two different time periods: one month before and after the TSP conclusion date (October 1, 2018), and another month before and after the beginning date of TSP (October 1, 2016). Thus, our non-pilot sample contains TAQ from September 1, 2016 to September 30, 2016 and from November 1, 2018 to November 30, 2018. The pilot sample includes TAQ from November 1 2016 to November 30 2016 and from September 1, 2018 to September 30, 2018. We follow Holden

[6] Congress passed the Jumpstart Our Business Startups Act ("Jobs Act") in 2012 with the goal of increasing the number of initial public offerings (IPOs) in the US markets with the idea that increased access to capital would lead to job creation by the smaller companies. The Jobs Act directed the SEC to conduct a study on how decimalization impacted the number of IPOs as well as the liquidity and trading of small-capitalization company stocks. The SEC decided to conduct a randomized trial to assess the impact of higher tick sizes on small firm stock liquidity, which can be particularly important for small firms as a number of papers including Amihud and Mendelson (1986), Brennan, Chordia, and Subrahmanyam (1998), and Brennan, Chordia, Subrahmanyam, and Tong (2012) have shown that higher liquidity leads to a lower cost of capital.

and Jacobsen (2014) in cleaning and matching the TAQ data. We drop all stocks without daily TAQ data, and transactions with negative or zero quoted spreads from the sample. We also filter out the stocks that are removed by the SEC from the test or control groups due to various reasons such as a price decline below \$1. Our final sample has 1007 (899), 391 (331), 383 (311) and 388 (317) stocks in the control, G_1, G_2, and G_3 groups, respectively over the time period September 1-30 and November 1-30, 2016 (2018) .

B. Regression Specification

To assess how the TSP impacted various liquidity measures, we implement the following difference-in-differences regression specification, consistent with the existing studies on the TSP (e.g., Rindi and Werner (2017) and Albuquerque et al. (2020)).

$$\begin{aligned} MQ_{it} =& \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event+ \\ & \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}, \end{aligned} \tag{16}$$

where MQ_{it} is a market quality or liquidity measure (e.g., quoted spreads, true spreads, effective spreads, etc.) for stock i at time t (millisecond-level); $Year$ is a time fixed-effect dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; *Event* is a dummy that equals one for all values during the pilot period; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and firm size that are available at beginning of trade t's transaction day. We cluster standard errors by firm and day. The coefficients β_5, β_6, and β_7 quantify the impact of the TSP on the liquidity of the treated groups G_1, G_2, and G_3, respectively. To conserve space, we only report the difference-in-differences coefficients β_5, β_6, and β_7 in the empirical section. We present tables containing all the regression coefficients in the Internet Appendix.

C. Testing the first hypothesis: Whether the TSP increased marking-making profits

1. Realized Spreads

To test whether the TSP improved market-making profits, we first assess two estimates of realized spreads: i) realized spreads that are computed using mid-quotes as proxies for the unknown true prices, and ii) realized spreads with this paper's rounding-adjusted prices as proxies for true prices.

Table 2 presents the difference-in-differences coefficients for the two realized spread measures. We find that the TSP significantly increased the rounding-adjusted realized spreads across all the treated stocks except for the G_1 unconstrained stocks. This result contrasts with the existing empirical studies (Rindi and Werner (2017) and Chung et al. (2019)) that use the mid-quote realized spreads to infer that the TSP did not increase market-making profits for providing liquidity to any of the unconstrained stocks. The first column in Table 2 shows that the increments using the mid-quote realized spreads across all the unconstrained treated stocks are economically and statistically insignificant.

Importantly, the increments to the rounding-adjusted realized spreads are consistent with the trading restrictions imposed by the TSP. Because investors can trade G_1 stocks (G_2 stocks) in 1 cent increments in all cases (some cases), the TSP is expected to increase market-making profits for providing liquidity to stocks in G_1, G_2, and G_3 in the order of $G_1 < G_2 < G_3$. Table 2 affirms that our realized spread increments align with this order across all samples, including the constrained and unconstrained. For example, the TSP increases the rounding-adjusted realized spreads of G_1 (G_1-constrained, G_1-unconstrained) by 0.8 (1, 0.2) cents, G_2 (G_2-constrained, G_2-unconstrained) by 1 (1.1, 0.9) cent, and G_3 (G_3-constrained, G_3-unconstrained) by 1.1 (1.2, 1.0) cents, with $0.8 < 1 < 1.1$ ($1 < 1.1 < 1.2$, $0.2 < 0.9 < 1$). Using bootstrap tests, we also find that this ordering pattern is statistically significant. In contrast, the mid-quote realized spreads do not satisfy the $G_1 < G_2 < G_3$ pattern. Thus, by capturing the key trading restriction imposed by the TSP, we validate that our rounding-adjusted inferences are more accurate than the existing mid-quote realized spreads.

2. Alternative Measures of Market-maker Profits: *mmp* and Realized Profits

Since the 1-minute realized spreads may not precisely capture true market-making profits given the surge in HFT (Conrad and Wahal (2020)), we also assess two other profit measures: *mmp* and realized profits; *mmp* equals the quoted spreads minus the true spreads and it captures hypothetical rounding profits earned by market-makers from a round-trip trade. However, for a given trade, market-makers earn rounding profits either from the ask or the bid quote, depending on whether the trade is a buy or a sell, but not both. Thus, we construct another measure, "realized profit", that equals transaction price minus the true ask if the trade is a customer buy (market-maker sell) and the true bid minus the transaction price if the trade is a customer sell (market-maker buy).

Table 3 presents the difference-in-differences coefficients of *mmp*, and realized profits. We find that the TSP signficantly increases the *mmp* and realized profits across all the treated stocks. The TSP significantly increased the *mmp* of G_1 (G_1-constrained, G_1-unconstrained) stocks by 3.5 (3.7, 3.1) cents, G_2 (G_2-constrained, G_2-unconstrained) stocks by 3.7 (3.9, 3.5) cents, and G_3 (G_3-constrained, G_3-unconstrained) stocks by 4.1 (4.1, 4.1) cents. Similarly, the TSP increased the realized profits of G_1 (G_1-constrained, G_1-unconstrained) stocks by 1.4 (1.4, 1.4) cents, G_2 (G_2-constrained, G_2-unconstrained) stocks by 1.6 (1.6, 1.8) cents, and G_3 (G_3-constrained, G_3-unconstrained) stocks by 1.9 (1.8, 2.2) cents. In the constrained stocks, the *mmp* and the realized profits increase mainly due to an increase in the quoted spreads, while in the unconstrained stocks the increase is mainly due to a decrease in the true spreads. In either case, the profits to market making due to rounding are economically and statistically significant.

Note that, the increments of both *mmp* and realized profits are ordered $G_1 < G_2 < G_3$, consistent with the trading restrictions imposed by the SEC. Also, the increments to the realized profits are significantly higher than those of the 1-minute realized spreads. This result, consistent with Conrad and Wahal (2020), who emphasize why and by how much the 1-minute realized spreads could underestimate true HFT market-making profits, as they quickly dissipate in a few seconds in the presence of HFTs.

Although the TSP increased the market-maker profits *per trade*, it could potentially decrease the aggregate market-maker profits, which equals per trade profits times the number of trades, if the number of trades are comparatively low during the pilot period. Thus, to address this question, we also assess the difference-in-differences of the share-weighted *mmp* and the share-weighted realized profits (rather than average *mmp* and average realized profits).[7] Table 4 presents the share-weighted difference-in-differences coefficients of *mmp* and realized profits. Note that the estimates of of the average *mmp* and the share weighted *mmp* are quite similar. The TSP significantly increases the market-maker profits across all the treated stocks, including the unconstrained stocks.

Interestingly, the magnitudes of realized profits suggest that sophisticated traders are able trace out true stock prices and liquidity. Tables 3 and 4 show that realized profits are generally less than half the *mmp*. This result suggests that, on average, liquidity demanders trade more at the quoted ask when the difference between the quoted and the true ask is lower than the difference between the true and the quoted bid. Similarly, liquidity demanders trade more at the quoted bid when the difference between the quoted and the true ask is higher than the difference between the true and the quoted bid. Thus, supplying liquidity is not as profitable as expected by naively comparing quoted and true spreads because the liquidity demanders are able to trade at prices where the impact of rounding is the lowest. Our results are consistent with Hagstromer (2020) and Muravyev and Pearson (2020), who argue that sophisticated traders use algorithms that allow them to ascertain the true cost of trading. This also motivates our use of sophisticated big-data methods to estimate true spreads and fundamental prices.

Overall, based on the rounding-adjusted measures, the TSP increased the market-maker profits across all the treated stocks, consistent with the SEC's first hypothesis.

3. Quoted Spreads and True Spreads

We now examine why our rounding-adjusted inferences differ from the existing inferences by examining the quoted spreads, the true spreads, and the components of the true spreads. Table 3 (4) presents the average (share-weighted) difference-in-differences coefficients of the quoted spreads

[7] Allena, Chen, and Chordia (2021) show that the share-weighted difference-in-differences could be estimated by minimizing the share-weighted least squares (rather than ordinary least squares) of the same regression as in (16). They also discuss how to obtain the standard errors of these share-weighted coefficients.

and true spreads. The TSP increased the quoted spreads (which measures market-makers' revenue per round-trip trade) across all the treated stocks except for the G_1 and G_2 unconstrained stocks in Table 3 and all the unconstrained stocks in Table 4. This result is not surprising, as the tick size does not bind for the unconstrained stocks. However, the TSP significantly decreased the true spreads (which measures the per round-trip trade costs incurred by the market-makers) across all the treated stocks, more so for the unconstrained stocks. As a result, market makers earn profits due to rounding for providing liquidity in the unconstrained stocks despite little or no change in their quoted spreads. The existing market-making profit measures, such as mid-quote realized spreads, ignore these TSP-driven decreases in true spreads, leading to biased conclusions that the TSP did not increase the market-maker profits for the unconstrained stocks.

Why did the TSP cause a decrease in the true spreads? To understand this, we examine the components of the spreads, viz., inventory and adverse selection costs.

D. Components of True Spreads

Our framework allows us to separately identify adverse selection costs, and the sum of order processing and inventory costs. Since order-processing costs such as computer costs, labor costs and informational service costs are largely fixed, we assume that the TSP does not impact these costs. As a result, we attribute the impact of the TSP on the true spreads to changes in the adverse selection and inventory components of the spread. The adverse selection and inventory component given the true spread s_t, are λs_t and $(1 - \lambda)s_t$, respectively. Table 5 shows that the TSP reduced both the adverse selection costs and the inventory costs across all treated stocks, and the declines are larger for the unconstrained stocks.

The decreases in the adverse selection and the inventory cost components of the true spreads are consistent with Goettler et al. (2005, 2009), who model dynamic limit order markets and explicitly account for rounding. In the case of constrained stocks, the true spreads are much lower than the quoted spreads due to the large binding tick size. The rounding driven artificially large quoted spreads reduce the adverse selection and inventory risk faced by market makers as the fundamental price is more likely to remain within the quoted spreads. This lowers the adverse selection and inventory risk components of the true spread. To extract the large rounding profits, the HFTs, given their speed advantage, provide liquidity by submitting fast limit orders to position themselves in the front of the limit order queues. These fast orders crowd out the relatively slow informed traders, who incorporate new information into stock prices. Although the informed traders could potentially undercut limit orders of HFTs, it is expensive due to the large tick size and often not possible as the minimum tick is binding. Thus, queue competition rather than price competition prevails in the tick-constrained environment resulting in larger depths at the quotes (O'Hara et al. (2019)), which we will document later. This queueing equilibirum is consistent with the model of Li et al. (2020)[8] as well as the results in Yao and Ye (2018). The queueing equilibrium leads to one more reason for why the adverse selection and liquidity risk are lower for the HFT market makers. As long as a particular HFT is not the first in the limit order queue, the other limit orders provide some buffer against the risk of being picked off.

In the case of the unconstrained stocks, undercutting of the resting limit orders is possible. As a result, due to the increased tick size informed traders strategically cut resting limit orders when it is profitable for them to do so. Since it is the informed traders, and not the market makers, who are likely to provide liquidity at the inside quotes, they are less likely to be subject to inventory risk as they do not carry inventory. The uninformed market makers have a lower incentive to provide the inside quotes because of the strategic undercutting by the informed traders, which would adversely select the uninformed market makers. Further, given their information advantage, the informed traders are also less subject to adverse selection. This is the undercutting equilibrium proposed by Goettler et al. (2005, 2009) in a limit order market with informed traders and it predicts lower adverse selection and inventory risk costs for the informed liquidity providers, which is what we find.[9]

[8] It is important to emphasize that Li et al. (2020) do not explicitly model informed traders, whereas our structural model does. However, our empirical results align with their theoretical predictions. The reason is that HFTs, because of their speed advantage, act on new information much faster than other traders, and thus they could be interpreted as informed traders (Biais, Foucault, and Moinas (2015)).

[9] Support for the undercutting equilibrium is provided by O'Hara et al. (2019), who use a unique NYSE data containing the order flows of institutional investors, quantitative traders, individual traders, and HFT market-makers, to document that a large tick size encourages informed HFTs to strategically undercut in the tick-unconstrained environment.

E. Validating our True Spread Components

We now validate our spread components using the aggregate market-maker participation data made publicly available by the SEC. We will show that the inventory component of our true spreads capture the TSP-driven decreased market-makers' inventory holdings, whereas the existing inventory cost measures of Huang and Stoll (1997) do not.

In particular, Appendix B (ii) of the SEC data contains the daily cumulative number of share buys and sells by all registered market-makers. Based on this daily trading activity of market-makers, we compute two measures that reflect the daily inventory costs/risks borne by them. The first measure is the absolute order imbalance, given by

$$Inv_{1it} = |\text{Number of Shares } Bought_{it} \text{ - Number of Shares } Sold_{it}|, \text{ for stock } i, \text{ on day } t. \quad (17)$$

Higher value of Inv_{1it} implies higher trade imbalance, and thus higher inventory costs/risk borne by the market-makers for stock i at the end of day t (Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes (2010)). Further, Chordia and Subrahmanyam (2004) argue that higher order imbalance may also reflect higher adverse selection risk when informed traders optimally choose to split their orders. However, they also assert that the order imbalances are significantly predicted by the lagged order imbalances, and this predictable component captures only the inventory risk but not the adverse selection risk. The intuition is that if today's high order imbalance predicts high order imbalance even for the next day, then it indicates high inventory holidng costs or risks borne by the market-makers. Recognizing this insight, Muravyev (2016) uses the order imbalance component that is predicted by the lagged order imbalance as measure of inventory holding risk. We use the same metric as another measure of inventory holding costs, given by

$$\begin{aligned}
&Inv_{2it} = |\hat{\alpha}_i + \hat{\beta}_{1i} OIB_{i,t-1} + \hat{\beta}_{2i} OIB_{i,t-2}, | \quad \text{where,} \\
&OIB_{it} = \alpha_i + \beta_{1i} OIB_{i,t-1} + \beta_{2i} OIB_{i,t-2} + \epsilon_{it}, \quad \text{and} \qquad (18) \\
&OIB_{it} = \text{Number of Shares } Bought_{it} \text{ - Number of Shares } Sold_{it}.
\end{aligned}$$

Table 6 shows that the difference-in-differences estimator of the two inventory risk measures are significantly negative across all but the G_1-constrained stocks, thus indicating that the TSP decreased the inventory costs borne by the market-makers. This result aligns with the TSP-driven decrease in the inventory component of our true spreads, as documented in Table 5.

Table 5 also shows the estimated adverse selection and the inventory cost components based on Huang and Stoll (1997) that do not account for rounding. Both of these components increase for the constrained stocks and remain unchanged for the unconstrained stocks. This result does not align with the market-maker participation data, which strongly indicates a significant decrease in the inventory costs across all the treated stocks. Thus, we validate that our rounding-adjusted spread components are more accurate than the existing spread measures that do not account for rounding.

F. Impact on Effective Spreads and Price Discovery

In the previous subsections, we showed that the TSP significantly increased the market-maker profits, consistent with the SEC's first hypothesis. This subsection shows that the TSP leads to a decline in liquidity as proxied by the effective spread, which is inconsistent with the SEC's second hypothesis.

1. Effective Spreads

We study the impact of the TSP on investor transaction costs by examining effective spreads, which are defined as twice the absolute value of the transaction price less a reference price,

$$Effective\ Spread_t = 2|P_t - P_t^R|, \tag{19}$$

where P_t is the transaction price at time t and P_t^R is the reference price. We use three different reference prices: (i) the mid-point of the bid-ask quote, $P_t^R = (A_t + B_t)/2$, (ii) the depth-weighted mid-point, $P_t^R = (B_t D_t^A + A_t D_t^B)/(D_t^A + D_t^B)$, and (iii) our rounding-adjusted true price, $P_t^R = m_t$. The depth-weighted reference price was proposed by Hagstromer (2020) who argues that the effective spreads that are computed using the depth-weighted reference prices adjust for the rounding biases, assuming all informed traders submit market orders.

Table 7 presents the difference-in-differences coefficients of the three effective spread proxies. Across all the treated stocks, but for the G_1 unconstrained stocks, the TSP significantly increased the effective spreads and thus the transaction costs for investors. Since the effective spread is twice the absolute difference between the transaction price and true price, it is not surprising that the increase in the tick size increased the effective spread. For the constrained stocks with a spread before the TSP of two cents it is clear that an increase in the tick size will increase the transaction price due to rounding and hence will also increase the effective spread. In the case of the unconstrained stocks, consider a tick grid of five cents starting at \$20. So if the inside bid (ask) prices are \$20.02 (\$20.12), then the requirement to quote on the grid will make the inside quoted bid (ask) equal \$20.00 (\$20.15). If the transaction were to take place at the quoted prices, then this would increase the effective spread.

In contrast, the mid-quote effective spreads, computed as in Rindi and Werner (2017), Albuquerque et al. (2020), and Chung et al. (2019), suggest that the TSP did not increase the transaction costs across all the unconstrained stocks. Thus, not accounting for rounding would lead to different inferences. In addition, we find that the depth-weighted effective spreads proposed by Hagstromer (2020) also indicate that the TSP did not impact the transaction costs of the unconstrained stocks, in contrast to this paper's results.

We validate that our effective spreads are more accurate measures of transaction costs than the mid-quote effective spreads and the depth-weighted effective spreads. During the TSP, investors can trade G_1 stocks in one cent increments in all cases and G_2 stocks in one cent increments in certain cases. Thus, the impact of TSP on *per trade* trading costs of investors must be in the order of $G_1 < G_2 < G_3$. Table 7 shows that our rounding-adjusted effective spreads precisely capture this pattern, whereas the mid-quote and Hagstromer's depth-weighted effective spreads do not.

The reason why Hagstromer's assertion that the depth-weighted effective spreads would effectively adjust the rounding biases need not hold, is the following. First, Hagstromer argued that the mid-quote effective spreads would *always* overstate the true effective spreads. On the contrary, we find that, during the non-pilot and the pilot periods, our rounding-adjusted effective spreads are generally higher than the mid-quote effective spreads across all stocks. In fact, even the depth-weighted effective spreads, on average, are higher than the mid-quote effective spreads. Both of these results do not align with Hagstromer (2020). We argue that this difference in results is likely driven by Hagstromer's assumption that informed liquidity demanders submit market orders or marketable limit orders. However, if the informed traders submit limit orders that are successfully executed, it is not necessarily the case that the mid-quote effective spreads would overstate the true effective spreads. Tables 5 and 9 provide evidence in support of this result, as informed traders switch from market orders to limit orders during the TSP.

Although the TSP increased the effective spreads *per trade*, it could potentially decrease the aggregate transaction costs of investors, proxied by effective spreads times the number of trades, if the number of trades are comparatively low during the pilot period. Thus, to examine the impact of the TSP on investors' aggregate transaction costs, table 8 presents the share-weighted difference-in-differences coefficients of the three effective spread measures. The conclusions are the same. Our rounding-adjusted effective spreads suggest that the TSP increased the transactions costs across all the treated stocks, including the unconstrained stocks. However, the other two effective spreads measures indicate that the TSP did not increase trasaction costs for the unconstrained stocks.

In sum, the increased market-maker profits due to the TSP have not led to a decrease in transaction costs as envisioned by the SEC.

2. Price Discovery

Our framework allows us to estimate the contribution of market, limit orders and new information to the price discovery. From equation (2), for each stock, we estimate the proportion of price discovery through market orders (R^2_M) as the R^2 in the regression of fundamental price changes on the signed half spread,

$$m_t - m_{t-1} = \lambda(q_t s_t)/2 + u_t. \tag{20}$$

Similarly, we estimate the proportion of price discovery through limit orders (R_L^2) as the R^2 in the regression of fundamental price changes on the changes in limit order book,

$$\begin{aligned} m_t - m_{t-1} &= \lambda_1(\Delta A_t)\,\mathrm{D}_t^A + \lambda_2(\Delta B_t)\,\mathrm{D}_t^B \\ &+\lambda_3\,(\Delta D_t^A)\,I_{\Delta A=0} + \lambda_4(\Delta D_t^B)\,I_{\Delta B=0} + v_t. \end{aligned} \tag{21}$$

Lastly, we estimate the proportion of price discovery through new information (R_I^2) as one minus the R^2 in the regression of fundamental price changes on the signed half spread and changes in the limit order book,

$$\begin{aligned} m_t - m_{t-1} &= \lambda(q_t s_t)/2 + \lambda_1(\Delta A_t)\,\mathrm{D}_t^A + \lambda_2(\Delta B_t)\,\mathrm{D}_t^B \\ &+\lambda_3\,(\Delta D_t^A)\,I_{\Delta A=0} + \lambda_4(\Delta D_t^B)\,I_{\Delta B=0} + \epsilon_t. \end{aligned} \tag{22}$$

The components of price discovery from market and limit orders could be correlated and thus it could be the case that $R_M^2 + R_L^2 + R_I^2 \neq 100\%$. To interpret the relative importance of the contributions price discovery through each channel and make uniform comparisons across various stocks and over different periods, we normalize the obtained R^2s so that $R_M^2 + R_L^2 + R_I^2 = 100\%$.

Table 9 presents the difference-in-differences coefficients of the proportions of price discovery through market orders, limit orders and new information. We find that the proportion of price discovery through limit orders increases and the proportions of price discovery through market orders and through new information decrease. This does not mean that the amount of new information in the economy had declined during the pilot period. What it means is that, relative to the control stocks, the TSP has led to a decrease in the *proportion* of price discovery through new information for the treated stocks. Given that a lower proportion of price discovery is through new information, prices are more responsive to previous trades and quotes rather than to new information, and thus are less efficient at least for the constrained stocks, as suggested by Hendershott, Jones, and Menkveld (2011) and Chordia, Green, and Kottimukkalur (2018).

It is not surprising to find a decline in price efficiency in the constrained stocks as the informed traders are crowded out in the queueing equilibrium. Due to the undercutting equilibrium in the unconstrained stocks, informed traders have more of an incentive to obtain fundamental information, which they incorporate into prices through limit orders, thereby decreasing the proportion of price discovery through new information.

3. Speed of Price Discovery

To evaluate whether the TSP increased the speed of price of discovery, we construct price delay measures of Chordia and Swaminathan (2000) and Hou and Moskowitz (2005) (that are also used in Chung et al. (2019) and Albuquerque et al. (2020)), but estimated using our rounding-adjusted true prices rather than the traditional mid-quote prices. We construct three delay measures D_1, D_2 and D_3 using the following regression specifications:

$$r_{it} = \alpha_{ic} + \beta_{ic} r_{mkt,t} + \eta_{cit}, \tag{23}$$

$$r_{it} = \alpha_i + \sum_{k=0}^{5} \beta_{ik} r_{mkt,t-k} + \eta_{it}, \tag{24}$$

where r_{it} denotes returns of stock i at time t computed with fundamental prices and $r_{mkt,t}$ is the return of the SPY index at time t. Consistent with earlier work, we aggregate returns over one minute time intervals. Equation (23) is the constrained regression of stock returns on the contemporaneous market returns, whereas equation (24) represents the unconstrained regression of stock returns on the contemporaneous and several lagged market returns. By construction, R^2 in the constrained regression, (R_c^2), is always less than or equal to the R^2 in the unconstrained regression, (R_u^2). If the stock i responds immediately to market returns, then β_{ic} significantly differs from zero but none of β_{ik}, $k > 0$ significantly differ from zero, and additionally $R_c^2 = R_u^2$. If, however, stock i responds with a lag then β_{ik} differ from zero and $R_c^2 < R_u^2$. Using the above insight, we have the following three metrics that quantify delays in price discovery:

$$
\begin{aligned}
D_1 &= 1 - \frac{R_c^2}{R_u^2} \\
D_2 &= \frac{\sum_{k=1}^{5} k|\beta_{ik}|}{\beta_{ic} + \sum_{k=1}^{5} k|\beta_{ik}|}, \\
D_3 &= \frac{\sum_{k=1}^{5} k|z_{ik}|}{z_{ic} + \sum_{k=1}^{5} k|z_{ik}|},
\end{aligned} \tag{25}
$$

where z_{ik} is the standard z-statistic for the coefficient estimate β_{ik}. A larger D_i implies that a stock's return responds with a delay to market returns. For each stock on a given day, we conduct independent regressions of equations (23) and (24) and obtain a panel data of all the above delay measures that span across all stocks and days in the sample.

Table 10 presents the difference-in-differences coefficients of the three delay measures. We find that the TSP did not improve the speed of price discovery across all the treated stocks. In fact, the TSP significantly delayed the price discovery across all the treated groups but the unconstrained. This result is consistent with Albuquerque et al. (2020), who also document that the TSP causes delays in price discovery, albeit using the mid-quotes as proxies for the unknown true prices.

Overall, the TSP significantly deteriorates liquidity in terms of increasing the transaction costs as proxied by the effective spreads. The TSP also reduces price efficiency by decreasing price discovery through new information, and by delaying price discovery (for the constrained stocks).

G. *Support for Theory*

We now examine market depths to provide support for the queue competition in constrained stocks and the undercutting equilibrium in the unconstrained stocks. Depth is the sum of the total bid and ask number of shares offered for trade at the inside bid and ask quotes. Table 11 presents the average depths conditional either on the quoted spreads or *mmp*.

It is clear that the average depths are far higher in the constrained stocks than in the unconstrained stocks as suggested by O'Hara et al. (2019). This is driven by queue competition where the HFTs compete to post quotes when the tick size is binding. Moreover, market-makers provide more depth during the periods of high profitability as proxied by *mmp*.[10] This suggests that market-makers (who are mainly the sophisticated HFTs) understand the true price and spread process, allowing them to provide more depth when it is profitable to do so. The sophisticated algorithms used by market makers, once again, motivates our use of the big data methodology in this paper.

We also provide evidence consistent with the TSP promoting strategic undercutting of informed traders in the unconstrained stocks. Table 11 shows that the average depth decreases with *mmp*, consistent with informed traders undercutting the resting limit orders during these periods.

VI. Conclusions

Observed prices and quoted spreads do not correspond to fundamental prices and true spreads when stocks are traded at prices rounded to the grid determined by the minimum tick size. Thus, the traditional liquidity and price efficiency measures that are not adjusted for rounding would be biased. This paper presents a structural model for true prices and spreads, explicitly accounting for rounding in the current HFT environment.

Estimating true liquidity and price measures from our structural model poses two main challenges. First, the rounding specification destroys the Gaussian error specification, rendering existing methods inapplicable. Second, the exisiting estimation techniques do not scale to the massive millisecond-level TAQ data. Thus, we develop a novel big-data methodology, Variational Inference, that scales to the TAQ data and estimates rounding-adjusted true prices and liquidity.

We apply our method to evaluate the recently conducted tick-size pilot program (TSP) and find that the TSP increased market-maker profits but decreased liquidity across all the treated stocks. This result contrasts existing empirical studies but are consistent with recent theoretical studies. Using the market-maker participation data made publicly available by the SEC, we validate that our rounding-adjusted liquidity measures are more accurate than the existing measures. We also find that some market participants, such as HFTs that use sophisticated algorithms, are able to trace out true price and quote process. This further validates our use of sophisticated big-data

[10]Note that the pattern of depth does not hold when conditioning on quoted spreads because, as we argued earlier, quoted spreads are not a good proxy for market maker profitability.

methods to estimate true prices and fundamental prices.

A. Appendix

This section derives the equations for updating the optimal density of a parameter or a hidden variable given other variables and parameters. First, the joint likelihood of the observed data $\{Y_1, Y_2, \ldots Y_T\}$ and the hidden variables $\{z_1, z_2, \ldots, z_T\}$, given the parameters Θ is :

$$\begin{aligned} & f\left(Y_{t=1}^T, z_{t=1}^T|\Theta\right) = f\left(Y_{t=1}^T|z_{t=1}^T,\Theta\right) f(z_{t=1}^T|\Theta) \\ & \propto \Pi_{t=1}^T \exp\left(-\frac{1}{2}\left(z_{t+1}-\mu_{t+1}-A_{t+1}z_t\right)^T \Sigma^{-1}\left(z_{t+1}-\mu_{t+1}-A_{t+1}z_t\right)\right) \mathcal{I}\left(I_{1t} \leq Bz_t \leq I_{2t}\right), \quad (26) \end{aligned}$$

where $z_t = \begin{bmatrix} x_t \\ \gamma_t \end{bmatrix}$, $\mu_t = \begin{bmatrix} \sum_{i=1}^4 l_i L_{it} \\ \alpha + \sum_{i=1}^4 d_i D_{it} \end{bmatrix}$, $A_t = \begin{bmatrix} 1 & \lambda q_t/2 \\ 0 & \beta \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1/2 \\ 1 & -1/2 \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix}$, and $\mathcal{I}(.)$ is the indicator function that takes the value 1 if the condition in (.)holds, takes zero otherwise, I_{1t}, I_{2t} are the bounds described in equation (8).

Under the diffuse prior specification for the parameters with $P(\Theta) = P\left(\Sigma, \lambda, \alpha, \beta, \{l_i, d_i\}_{i=1}^4\right) \propto \Sigma^{-\frac{2+1}{2}}$, update equations for each parameter and hidden state variable are derived as below:

Proposition 2. *The optimal density of λ, $q^*(\lambda)$ given the optimal densities of all other parameters and hidden variables is given by:*

$$q^*(\lambda) \sim \mathcal{N}\left(\frac{E(\sum_t A_t^\lambda B_t^\lambda)}{E(\sum A_t^{\lambda^2})}, \frac{E(\sigma_\epsilon^2)}{E(\sum_t A_t^{\lambda^2})}\right), \quad (27)$$

where $A_t = \gamma_{t-1}$, $B_t = x_t - x_{t-1} - \sum_{i=1}^4 l_i L_{it}$, and the expectations are taken with respect to the optimal variational densities of parameters and state variables other than λ.

Proof.

$$q^*(\lambda) \propto \exp\left[-\sum_{t=1}^{T} \frac{E\left(x_t - x_{t-1} - \lambda\gamma_{t-1} - \sum_{i=1}^{4} l_i L_{it}\right)^2}{2E(\sigma_\epsilon^2)}\right]$$

$$\propto exp\left[-\sum_{t=1}^{T} \frac{\left(B_t^\lambda - \lambda A_t^\lambda\right)^2}{2E(\sigma_\epsilon^2)}\right] = \exp\left[-\frac{\left(\lambda - \frac{E(\sum_t A_t^\lambda B_t^\lambda)}{\sum_t A_t^{\lambda 2}}\right)^2}{2E(\sigma_\epsilon^2)/E(\sum_t A_t^{\lambda 2})}\right]$$

$$\implies q^*(\lambda) \sim \mathcal{N}\left(\frac{E(\sum_t A_t^\lambda B_t^\lambda)}{E(\sum A_t^{\lambda 2})}, \frac{E(\sigma_\epsilon^2)}{E(\sum_t A_t^{\lambda 2})}\right) \tag{28}$$

□

Proposition 3. *The optimal density of l_i, $q^*(l_i)$ given the optimal densities of all other parameters and hidden variables is given by:*

$$q^*(l_i) \sim \mathcal{N}\left(\frac{E\left(\sum_{t=1}^{T} A_t^{l_i} L_{it}\right)}{\sum_{t=1}^{T} L_{it}^2}, \frac{E(\sigma_\epsilon^2)}{\sum_{t=1}^{T} L_{it}^2}\right), \tag{29}$$

where $A_t^{l_i} = x_t - x_{t-1} - \sum_{j \neq i} l_j L_{jt}$, and the expectation is taken with respect to optimal variational densities of parameters and state variables other than l_i.

Proof.

$$q^*(l_i) \propto \exp\left[-\sum_{t=1}^{T} E\left(l_i L_{it} - A_t^{l_i}\right)^2 / 2E(\sigma_\epsilon^2)\right] \propto \exp\left[-\frac{\left(l_i - \frac{E(\sum_{t=1}^{T} A_t^{l_i} L_{it})}{\sum_{t=1}^{T} L_{it}}\right)^2}{2E(\sigma_\epsilon^2)/(\sum_{t=1}^{T} L_{it})^2}\right]$$

$$\implies q^*(\beta) \sim \mathcal{N}\left(\frac{E(\sum_{t=1}^{T} \gamma_{t-1} A_t^\gamma)}{E\left(\sum_{t=1}^{T} \gamma_{t-1}^2\right)}, \frac{E(\sigma_\eta^2)}{E(\sum_{t=1}^{T} \gamma_{t-1}^2)}\right) \tag{30}$$

□

Proposition 4. *The optimal density of β, $q^*(\beta)$ given the optimal densities of all other parameters and hidden variables is given by:*

$$q^*(\beta) \sim \mathcal{N}\left(\frac{E(\sum_{t=1}^{T}\gamma_{t-1}A_t^{\gamma})}{E\left(\sum_{t=1}^{T}\gamma_{t-1}^2\right)}, \frac{E(\sigma_\eta^2)}{E(\sum_{t=1}^{T}\gamma_{t-1}^2)}\right), \tag{31}$$

where $A_t^{\gamma} = \gamma_t - \alpha - (\sum_{i=1}^{4} d_i D_{it})$, and the expectations are taken with respect to the variational densities of all the parameters and variables, but β.

Proposition 5. *The optimal density of each d_i, $q^*(d_i)$ given the optimal densities of all other parameters and hidden variables is given by:*

$$q^*(d_i) \sim \mathcal{N}\left(\frac{\sum E(D_{it}A_t^{D_i})}{\sum_{t=1}^{T} D_{it}^2}, \frac{E(\sigma_\epsilon^2)}{\sum_{t=1}^{T} D_{it}^2}\right), \tag{32}$$

where $A_t^{D_i} = \gamma_t - \left(\alpha + \beta\gamma_{t-1} + \sum_{k\neq i} D_{kt}d_k\right)$

Proposition 6. *The optimal density of each state variable z_t, $q^*(z_t)$ given the optimal densities of all other parameters and hidden variables is given by:*

$$z_t \sim \mathcal{N}\left({\Sigma^*}^{-1}\left(\Sigma^{-1}\mu_{1t} + A_{t+1}^T\Sigma^{-1}\mu_{2t}\right), {\Sigma^*}^{-1}\right), I_{1t} \leq Bz_t \leq I_{2t}, \tag{33}$$

where $\mu_{1t} = A_tE(Z_{t-1}) + E(\mu_t)$ and $\mu_{2t} = E(z_{t+1})$; $\Sigma^* = \Sigma^{-1} + A_{t+1}^T\Sigma^{-1}A_{t+1}$; B, I_{1t}, I_{2t} are given in equation (26).

Proof.

$$\begin{aligned} q^*(z_t) \propto \exp&\left(-\frac{1}{2}E\left[\left(z_{t+1} - \mu_{t+1} - A_{t+1}z_t\right)^T\Sigma^{-1}\left(z_{t+1} - \mu_{t+1} - A_{t+1}z_t\right)\right]\right) \times \\ &\exp\left(-\frac{1}{2}E\left[\left(z_t - \mu_t - A_tz_{t-1}\right)^T\Sigma^{-1}\left(z_t - \mu_t - A_tz_{t-1}\right)\right]\right) \times \mathcal{I}(I_{1t} \leq Bz_t \leq I_{2t}) \end{aligned}$$

Letting $\mu_{1t} = A_t E(Z_{t-1}) + E(\mu_t)$ and $\mu_{2t} = E(z_{t+1})$, and $\Sigma^* = \Sigma^{-1} + A_{t+1}^T \Sigma^{-1} A_{t+1}$, we have

$$q^*(z_t) \propto \mathcal{I}(I_{1t} \leq A z_t \leq I_{2t}) \times$$
$$exp\left(-\frac{1}{2}\left(z_t - {\Sigma^*}^{-1}\left(\Sigma^{-1}\mu_{1t} + A_{t+1}^T\Sigma^{-1}\mu_t\right)\right)^T \Sigma^* \left(z_t - {\Sigma^*}^{-1}\left(\Sigma^{-1}\mu_{1t} + A_{t+1}^T\Sigma^{-1}\mu_t\right)\right)\right)$$

Therefore,

$$z_t \sim \mathcal{N}\left({\Sigma^*}^{-1}\left(\Sigma^{-1}\mu_{1t} + A_{t+1}^T\Sigma^{-1}\mu_t\right), {\Sigma^*}^{-1}\right), I_{1t} \leq B z_t \leq I_{2t} \tag{34}$$

□

Thus, the optimal density of z_t follows a linearly constrained (truncated) version of a normally distributed random variable, given the moments of all other parameters and the hidden variables, z_{t-1} and z_{t+1}. The first two moments of a linearly constrained normal random variable can be computed using the procedure outlined by Kan and Robotti (2017).

References

Albuquerque, Rui A., Shiyun Song, and Chen Yao, 2020, The price effects of liquidity shocks: A study of sec's tick-size experiment, *Journal of Financial Economics (JFE), Forthcoming* .

Allena, Rohit, 2021, Confident risk premiums and investments using machine learning uncertainties, *SSRN Electronic Journal* .

Allena, Rohit, Te-Feng Chen, and Tarun Chordia, 2021, Microstructure inferences with the millisecond-level taq data, *Working paper, Emory University* .

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Ball, Clifford A, and Tarun Chordia, 2001, True spreads and equilibrium prices, *The Journal of Finance* 56, 1801–1835.

Bessembinder, Hendrik, 2003, Issues in assessing trade execution costs, *Journal of Financial Markets* 6, 233–257.

Biais, Bruno, Thierry Foucault, and Sophie Moinas, 2015, Equilibrium fast trading, *Journal of Financial Economics* 116, 292–313.

Blei, David M., Alp Kucukelbir, and Jon D. McAuliffe, 2017, Variational inference: A review for statisticians, *Journal of the American Statistical Association* 112, 859–877.

Brennan, Michael J, Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 354–373.

Brennan, Michael J., Tarun Chordia, Avanidhar Subrahmanyam, and Qing Tong, 2012, Sell-order liquidity and the cross-section of expected stock returns, *Journal of Financial Economics* 105, 523–541.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2019, Price discovery without trading: Evidence from limit orders, *The Journal of Finance* 74, 1621–1658.

Chinco, Alex, Adam D. Clark-Joseph, and Mao Ye, 2019, Sparse signals in the cross-section of returns, *Journal of Finance* 74, 449–492.

Chordia, Tarun, T. Clifton Green, and Badrinath Kottimukkalur, 2018, Rent seeking by low-latency traders: Evidence from trading on macroeconomic announcements, *Review of Financial Studies* 31, 4650–4687.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2001, Market liquidity and trading activity, *The Journal of Finance* 56, 501–530.

Chordia, Tarun, and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485–518.

Chordia, Tarun, and Bhaskaran Swaminathan, 2000, Trading volume and cross-autocorrelations in stock returns.

Chung, Kee H., Albert Lee, and Dominik Roesch, 2019, Tick size, liquidity for small and large orders, and price informativeness: Evidence from the tick size pilot program, *Journal of Financial Economics (JFE), Forthcoming* .

Comerton-Forde, Carole, Vincent Grégoire, and Zhuo Zhong, 2019, Inverted fee structures, tick size, and market quality, *Journal of Financial Economics* .

Comerton-Forde, Carole, Terrence Hendershott, Charles M. Jones, Pamela C. Moulton, and Mark S. Seasholes, 2010, Time variation in liquidity: The role of market-maker inventories and revenues, *Journal of Finance* 65, 295–331.

Conrad, Jennifer, and Sunil Wahal, 2020, The term structure of liquidity provision, *Journal of Financial Economics* 136, 239–259.

Demsetz, Harold, 1968, The cost of transacting, *The Quarterly Journal of Economics* 82, 33.

Easley, David, and Maureen O'Hara, 1992, Time and the process of security price adjustment, *The Journal of Finance* 47, 577–605.

Glosten, Lawrence R., and Paul R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.

Goettler, Ronald L, Christine A Parlour, and Uday Rajan, 2005, Equilibrium in a dynamic limit order market, *The Journal of Finance* 60, 2149–2192.

Goettler, Ronald L., Christine A. Parlour, and Uday Rajan, 2009, Informed traders and limit order markets, *Journal of Financial Economics* 93, 67–87.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *The Review of Financial Studies* 33, 2223–2273.

Hagstromer, Bjorn, 2020, Bias in the effective bid-ask spread, *Journal of Financial Economics (JFE), Forthcoming* .

Hasbrouck, Joel, 1999a, The dynamics of discrete bid and ask quotes, *Journal of Finance* 54, 2109–2142.

Hasbrouck, Joel, 1999b, Security bid/ask dynamics with discreteness and clustering: Simple strategies for modeling and estimation, *Journal of Financial Markets* 2, 1–28.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld, 2011, Does algorithmic trading improve liquidity?, *Journal of Finance* 66, 1–33.

Holden, Craig W, and Stacey Jacobsen, 2014, Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions, *THE JOURNAL OF FINANCE •* LXIX.

Hou, Kewei, and Tobias J. Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns.

Huang, Roger D., and Hans R. Stoll, 1997, The components of the bid-ask spread: A general approach, *Review of Financial Studies* 10, 995–1034.

Kalman, R E, 1960, A new approach to linear filtering and prediction problems, *Transactions of the ASME-Journal of Basic Engineering* 82, 35–45.

Kan, Raymond, and Cesare Robotti, 2017, On moments of folded and truncated multivariate normal distributions, *Journal of Computational and Graphical Statistics* 26, 930–934.

Korajczyk, Robert A., and Ronnie Sadka, 2008, Pricing the commonality across alternative measures of liquidity, *Journal of Financial Economics* 87, 45–72.

Li, Sida, Xin Wang, and Mao Ye, 2020, Who provides liquidity, and when?, *Journal of Financial Economics (JFE), Forthcoming* .

Muravyev, Dmitriy, 2016, Order flow and expected option returns, *Journal of Finance* 71, 673–708.

Muravyev, Dmitriy, and Neil D. Pearson, 2020, Options trading costs are lower than you think, *The Review of Financial Studies* 33, 4973–5014.

O'Hara, Maureen, Gideon Saar, and Zhuo Zhong, 2019, Relative tick size and the trading environment, *The Review of Asset Pricing Studies* 9, 47–90.

Rindi, Barbara, and Ingrid M. Werner, 2017, U.s. tick size pilot, *SSRN Electronic Journal* .

Stoikov, Sasha, 2018, The micro-price: a high-frequency estimator of future prices, *Quantitative Finance* 18, 1959–1966.

Stoll, Hans R, 1978, The supply ofdealer services in securities markets, *The Journal of Finance* XXXl, 1133–1151.

Wang, Yixin, and David M. Blei, 2018, Frequentist consistency of variational bayes, *Journal of the American Statistical Association* 1–15.

Yao, Chen, and Mao Ye, 2018, Why trading speed matters: A tale of queue rationing under price controls, *The Review of Financial Studies* 31, 2157–2183.

Figure 1. ELBO (y-axis) vs Number of Epochs (x-axis)

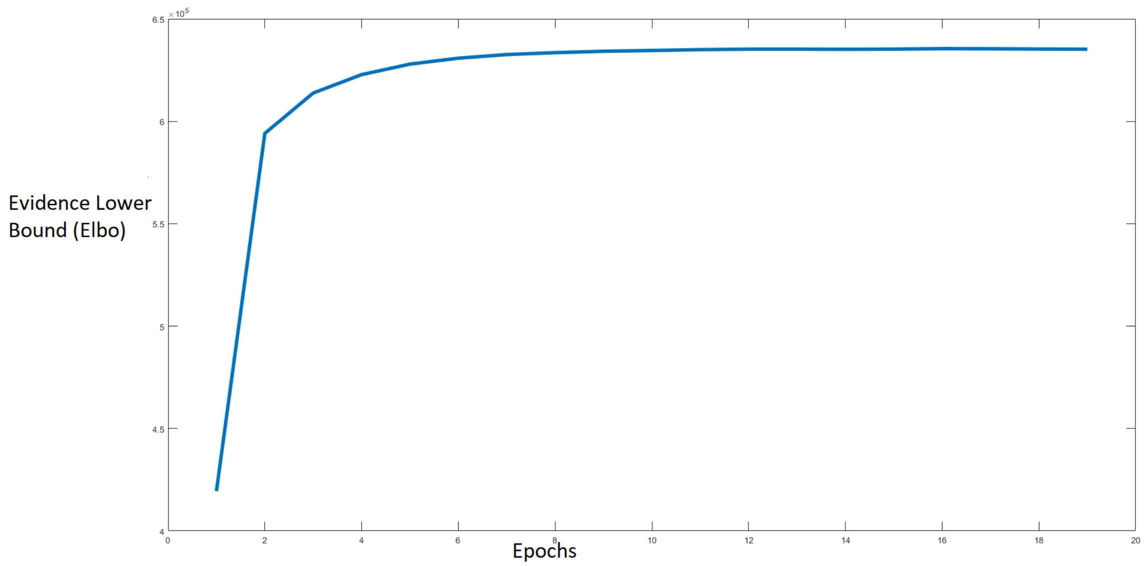

Table 1. Performance of the Methodology: Monte Carlo Evidence

This table reports the performance of our methodology in estimating the true simulated parameters. We conduct Monte Carlo simulations, where the moments of the simulated true prices and spreads are calibrated to match the moments of the mid-quotes and half times the bid-ask spreads, respectively, of three randomly selected stocks, FCB, SAM and AMC. The "True Sim Val" column represents the true simulated means and standard deviations of various liquidity measures, such as true spreads (tspr), true prices, true effective spreads (espr), market-maker profits per share traded (mmp), and the price impact. The "Estimated Val" column presents the estimated means and standard deviations of various liquidity measures using this paper's methodology. The "Naive Val" column depicts the naively estimated liquidity measurements using the observed transaction prices and quotes without adjusting for rounding. All of the above values are reported in dollars. The table also shows the squared sum of errors in estimating fundamental prices and true spreads using this paper's method (i.e., with the rounding adjustment) and the naive method (i.e., without the rounding adjustment).

Stock	Variable	True Sim Val	Est Val	Naïve Val
FCB	mean tspr	0.0441	0.0445	0.0941
	std tspr	0.0275	0.0215	0.0338
	mean price	58.3545	58.3538	58.3489
	std price	14.3335	14.3349	14.3326
	mean espr	0.0861	0.0859	0.0941
	std espr	0.0367	0.0365	0.0340
	mean mmp	0.0501	0.0497	0.0941
	std mmp	0.0199	0.0221	0.0338
	Price Impact (λ)	0.1923	0.1927	0.0917
	Estimation Error - tspr		15.8693	170.6501
	Estimation Error - price		9.3137	45.9682
SAM	mean tspr	0.1817	0.1821	0.2318
	std tspr	0.1221	0.1146	0.1237
	mean price	222.2710	222.2699	222.2585
	std price	38.1229	38.1425	38.1226
	mean espr	0.1931	0.1936	0.2318
	std espr	0.1006	0.1021	0.1237
	mean mmp	0.0501	0.0497	0.2318
	std mmp	0.0197	0.0269	0.1237
	Price Impact (λ)	0.1858	0.1888	0.1530
	Estimation Error - tspr		27.5806	126.3707
	Estimation Error - price		11.6936	536.0978
AMC	mean tspr	0.0459	0.0460	0.0960
	std tspr	0.0247	0.0278	0.0318
	mean price	37.1058	37.1046	37.1022
	std price	2.5691	2.5748	2.5687
	mean espr	0.0925	0.0926	0.0960
	std espr	0.0368	0.0368	0.0318
	mean mmp	0.0501	0.0500	0.0960
	std mmp	0.0201	0.0188	0.0318
	Price Impact (λ)	0.0709	0.0695	0.0353
	Estimation Error - tspr		9.6458	157.3844
	Estimation Error - price		8.1756	42.2229

Table 2. Rounding-adjusted Realized spreads and Mid-quote Realized Spreads

This table presents the difference-in-differences coefficient that quantify the average changes in liquidity measures due to the TSP program. The liquidity measures include two different estimates of realized spreads: 1) realized spreads computing using this paper's estimated true prices; and 2) realized spreads computed using mid-quotes as proxies for true prices. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure (e.g., realized spread) for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Mid-quote realized spreads		True price realized spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.007	5.236	0.008	5.025
G2× Event	0.008	8.885	0.010	6.765
G3 × Event	0.005	8.193	0.011	42.925

Constrained Stocks

	Mid-quote realized spreads		True price realized spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.010	7.560	0.010	5.590
G2× Event	0.011	56.598	0.011	5.589
G3 × Event	0.009	5.717	0.012	6.226

Unconstrained Stocks

	Mid-quote realized spreads		True price realized spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.000	-0.050	0.002	0.446
G2× Event	0.005	1.030	0.009	3.420
G3 × Event	0.001	0.807	0.010	2.279

Table 3. Quoted Spreads, True Spreads and Market-Maker Profits

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The liquidity measures include quoted spreads, true spreads, market-maker profits per share traded, and realized market-maker profits per-share traded. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure (e.g., quoted spread) for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.025	17.716	-0.012	-7.456	0.036	120.890	0.014	72.503
G2× Event	0.021	9.212	-0.017	-9.522	0.038	155.858	0.016	89.870
G3 × Event	0.026	13.311	-0.016	-8.267	0.042	110.453	0.019	77.874

Constrained Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.034	27.040	-0.004	-2.064	0.038	190.935	0.014	78.052
G2× Event	0.034	36.980	-0.005	-4.199	0.039	147.316	0.016	77.018
G3 × Event	0.035	37.513	-0.007	-4.937	0.042	155.173	0.018	86.848

Unconstrained Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.003	0.877	-0.030	-10.715	0.033	71.751	0.014	44.063
G2× Event	0.003	0.635	-0.034	-8.305	0.037	65.062	0.018	80.997
G3 × Event	0.007	2.052	-0.036	-7.817	0.043	53.080	0.022	50.739

Table 4. Quoted Spreads, True Spreads and Market-Maker Profits: Share-weighted
This table presents the share-weighted difference-in-differences coefficients (i.e., $\beta_5, \beta_6, \beta_7$) that quantify the average changes in liquidity measures due to the TSP program. The liquidity measure include quoted spreads, true spreads, market-maker profits per share traded, and realized market-maker profits per-share traded. The difference-in-differences coefficients minimize the share-weighted least squares (rather than ordinary least squares (OLS)) in the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$,
where MQ_{it} is a market quality measure (e.g., quoted spread) for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.026	20.426	-0.010	-3.845	0.035	80.009	0.012	85.647
G2× Event	0.023	13.396	-0.014	-7.087	0.037	112.763	0.015	85.292
G3 × Event	0.026	20.090	-0.015	-7.807	0.041	114.152	0.018	69.721

Constrained Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.034	28.725	-0.003	-2.674	0.037	108.943	0.013	56.076
G2× Event	0.034	27.260	-0.004	-3.023	0.039	63.154	0.015	71.467
G3 × Event	0.034	31.143	-0.007	-4.820	0.041	119.132	0.017	95.645

Unconstrained Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	-0.001	-0.265	-0.032	-7.346	0.031	50.076	0.012	45.825
G2× Event	0.000	-0.056	-0.035	-8.519	0.035	59.001	0.016	62.856
G3 × Event	0.004	0.908	-0.037	-7.383	0.041	39.125	0.021	22.269

Table 5. Components of Bid-Ask Spread

This table presents the difference-in-differences coefficients (i.e., $\beta_5, \beta_6, \beta_7$) that quantify the average changes in liquidity measures due to the TSP program. The liquidity measures include adverse selection and inventory components of spreads per share traded. These components are computed using two approaches: 1) using this paper's method, explicitly accounting for the rounding specification; 2) using Huang and Stoll (1997) method, without accounting for rounding. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure (e.g., adverse selection) for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

	Components with the rounding adjustment: All Stocks					Components without the rounding adjustment: All Stocks			
	Adverse selection		Inventory costs			Adverse selection		Inventory costs	
Coefficient	Estimate	t-stat	Estimate	t-stat	Coefficient	Estimate	t-stat	Estimate	t-stat
G1 × Event	-0.002	-12.572	-0.009	-13.047	G1 × Event	0.005	6.094	0.019	16.246
G2× Event	-0.002	-5.609	-0.014	-7.635	G2× Event	0.005	8.352	0.016	11.643
G3 × Event	-0.002	-5.606	-0.014	-11.843	G3 × Event	0.004	5.975	0.021	19.474

	Components with the rounding adjustment: Tick-constrained stocks					Components without the rounding adjustment: Tick-constrained stocks			
	Adverse selection		Inventory costs			Adverse selection		Inventory costs	
Coefficient	Estimate	t-stat	Estimate	t-stat	Coefficient	Estimate	t-stat	Estimate	t-stat
G1 × Event	-0.001	-2.901	-0.003	-2.054	G1 × Event	0.008	11.993	0.026	16.936
G2× Event	-0.001	-2.005	-0.005	-4.462	G2× Event	0.008	44.559	0.026	39.349
G3 × Event	-0.001	-3.782	-0.006	-5.985	G3 × Event	0.007	31.473	0.028	76.495

	Components with the rounding adjustment: Tick-unconstrained stocks					Components without the rounding adjustment: Tick-unconstrained stocks			
	Adverse selection		Inventory costs			Adverse selection		Inventory costs	
Coefficient	Estimate	t-stat	Estimate	t-stat	Coefficient	Estimate	t-stat	Estimate	t-stat
G1 × Event	-0.006	-9.720	-0.024	-10.182	G1 × Event	0.000	-0.410	0.003	1.287
G2× Event	-0.005	-5.141	-0.028	-9.152	G2× Event	0.001	1.027	0.001	0.543
G3 × Event	-0.004	-6.184	-0.032	-10.021	G3 × Event	-0.001	-0.863	0.008	2.032

Table 6. Inventory Risks of Aggregate Market-Makers
This table reports the averages of two measures of inventory costs estimated directly from the SEC market-makers' participation data, across all the treated groups G_1, G_2, G_3, and the Control group. The inventory costs are estimated using order imbalances, measured as the difference in the number of shares bought less the number of shares sold. Constrained (unconstrained) stocks are those whose quoted bid-ask spreads were lower (higher) than 5 cents prior to the TSP. The column "Non-Pilot" ("Pilot") presents the estimated mean values of variables of each group during the non-pilot (pilot) regime. The column "Diff" is the difference estimate of variables of each group, prior and post the pilot program. "DD" is the difference-in-differences estimate of variables in treated groups with respect to the variables in control group. Standard errors are in parenthesis and ** denotes significance at the 5% level.

Inventory Costs of Aggregate Market-Makers

Non-Pilot = Sep 1-30, 2016 & Nov 1-30, 2018 ; Pilot = Nov 1-30, 2016 & Sep 1-30, 2018

	Order Imbalance (Inv_{1it})				Expected Order Imbalance (Inv_{2it})			
Group	Non-Pilot	Pilot	Diff	DD	Non-Pilot	Pilot	Diff	DD
G1 Constrained	13375	17597	4221^{**} (493)	1897^{**} (378)	6587	8402	1815^{**} (200)	713^{**} (148)
G1 Unconstrained	2595	2732	137 (89)	-2187^{**} (391)	1614	1337	-278^{**} (98)	-1380^{**} (156)
G1	7321	9333	2012^{**} (203)	-312 (234)	3765	4536	772^{**} (84)	-1180^{**} (82)
G2 Constrained	11711	13796	2085^{**} (322)	-239 (382)	6266	6683	417^{**} (155)	-685^{**} (151)
G2 Unconstrained	2362	2149	-212^{**} (54)	-2536^{**} (389)	1202	1054	-149^{**} (29)	-1251^{**} (153)
G2	6640	8109	1468^{**} (146)	-855^{**} (224)	3501	3971	471^{**} (62)	-631^{**} (88)
G3 Constrained	11847	11257	-590^{**} (290)	-2914^{**} (359)	5806	5656	-150 (119)	-1252^{**} (140)
G3 Unconstrained	2615	2503	-112 (120)	-2436^{**} (401)	1503	1832	329 (329)	-773^{**} (184)
G3	6701	6432	-269^{**} (132)	-2593^{**} (222)	3417	3597	180 (126)	-923^{**} (108)
Control	7179	9502	2324^{**} (118)		3653	4755	1102^{**} (46)	

Table 7. Effective Spreads with Fundamental Prices, Mid-quotes, and Weighted Mid-quotes
This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The liquidity measures include effective spreads computed using three different approaches: 1) using this paper's method; 2) using mid quotes as proxies for true prices; 3) using Hagstromer (2020) method. The difference-in-differences coefficients minimize the share-weighted least squares (rather than ordinary least squares (OLS)) in the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure (e.g., effective spread) for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks						
	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.017	25.788	0.015	9.616	0.018	17.831
G2× Event	0.015	10.562	0.013	11.628	0.019	11.193
G3 × Event	0.015	11.401	0.013	8.983	0.024	14.923

Constrained Stocks						
	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.025	22.199	0.021	16.907	0.025	12.993
G2× Event	0.025	76.437	0.021	38.504	0.027	15.938
G3 × Event	0.023	117.106	0.020	92.987	0.030	31.777

Unconstrained Stocks						
	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.000	0.049	0.001	0.362	0.004	1.257
G2× Event	0.001	0.204	0.001	0.237	0.006	1.997
G3 × Event	-0.002	-0.650	-0.001	-0.391	0.013	4.308

Table 8. Effective Spreads: Share-Weighted

This table presents the difference-in-difference coefficients that quantify the average changes in liquidity measures due to the TSP program. The liquidity measures include effective spreads computed using three different approaches: 1) using this paper's method; 2) using mid quotes as proxies for true prices; 3) using Hagstromer (2020) method. The difference-in-difference coefficients minimize the share-weighted least squares (rather than ordinary least squares (OLS)) in the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure (e.g., effective spread) for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.019	14.109	0.016	14.237	0.019	18.820
G2× Event	0.017	10.687	0.015	12.127	0.020	14.400
G3 × Event	0.017	12.914	0.014	14.230	0.025	20.443

Constrained Stocks

	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.024	29.638	0.020	13.762	0.024	18.716
G2× Event	0.025	31.139	0.021	20.032	0.027	29.617
G3 × Event	0.023	59.785	0.020	21.111	0.030	34.493

Unconstrained Stocks

	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.000	-0.101	0.001	0.179	0.003	0.835
G2× Event	0.001	0.288	0.001	0.340	0.006	2.359
G3 × Event	-0.003	-0.560	-0.002	-0.332	0.012	2.410

Table 9. Proportions of Price Discovery

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The liquidity measures include proportions of price discovery through market orders, limit orders, and new information. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$,

where MQ_{it} is a market quality measure (e.g., market order price discovery) for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Market Order Price Discovery		Limit Order Price Discovery		New Info Price Discovery	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	-0.043	-8.399	0.033	6.112	-0.010	-16.292
G2× Event	-0.047	-14.586	0.038	16.095	-0.008	-12.975
G3 × Event	-0.040	-14.868	0.032	9.695	-0.008	-14.860

Constrained Stocks

	Market Order Price Discovery		Limit Order Price Discovery		New Info Price Discovery	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	-0.043	-5.998	0.032	5.293	-0.011	-17.047
G2× Event	-0.051	-16.103	0.041	20.097	-0.010	-10.664
G3 × Event	-0.046	-9.499	0.037	9.846	-0.010	-15.409

Unconstrained Stocks

	Market Order Price Discovery		Limit Order Price Discovery		New Info Price Discovery	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	-0.044	-16.124	0.036	16.940	-0.008	-10.842
G2× Event	-0.041	-8.459	0.034	8.987	-0.006	-9.707
G3 × Event	-0.027	-6.149	0.022	5.177	-0.005	-4.646

Table 10. Speed of Price Discovery

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The liquidity measures include delay measures $D1$, $D2$, $D3$ of Chordia and Swaminathan (2000) and Hou and Moskowitz (2005) that signify the speeds of price discovery. The higher the delay measure is, the slower the price discovery will be. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure (e.g., delay measure D1) for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Delay Measure D1		Delay Measure D2		Delay Measure D3	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.087	8.294	0.037	4.272	0.037	4.185
G2× Event	0.071	5.740	0.029	4.455	0.029	3.571
G3 × Event	0.092	7.524	0.037	8.380	0.037	5.981

Constrained Stocks

	Delay Measure D1		Delay Measure D2		Delay Measure D3	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	0.129	5.913	0.058	5.545	0.058	5.698
G2× Event	0.122	7.139	0.049	7.648	0.050	8.088
G3 × Event	0.133	8.783	0.057	7.997	0.057	8.822

Unconstrained Stocks

	Delay Measure D1		Delay Measure D2		Delay Measure D3	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
G1 × Event	-0.019	-1.220	-0.014	-1.892	-0.015	-1.726
G2× Event	-0.019	-1.577	-0.007	-1.013	-0.008	-1.249
G3 × Event	-0.001	-0.075	-0.009	-1.258	-0.009	-1.040

Table 11. Market-maker Profits and Depths

This table reports the average estimated total bid-ask depth at the best bid and best ask quotes conditional on market maker profits. For every stock on each day, we split the transactions into two depending on their profitability. "High_{mmp}" ("Low_{mmp}") represents the subsample of transactions that have above (below) the median-market maker profits. "Average depth" column in Panel A represents the average of the total depth (sum of the best bid and best ask depths) on the "High_{mmp}" and "Low_{mmp}" subsamples across all stocks in various groups, such as G1, G2, and G3. "High_{qspr}" ("Low_{qspr}") represents the subsample of transactions that have above (below) the quoted spreads for each stock, each trading day. "Average depth" column in Panel B is the average of the total depth on the "High_{qspr}" and "Low_{qspr}" subsamples. Constrained (unconstrained) stocks are those whose quoted bid-ask spreads were lower (higher) than 5 cents prior to the TSP. The sample period is from September 1, 2018 to September 30, 2018.

	Panel A		Panel B	
Group	High/Low *mmp*	Average Depth	High/Low Quoted Spread	Average Depth
G1 Constrained	High_{mmp}	7638.53	High_{qspr}	6580.05
	Low_{mmp}	6232.22	Low_{qspr}	7100.35
G1 Unconstrained	High_{mmp}	815.40	High_{qspr}	736.97
	Low_{mmp}	837.44	Low_{qspr}	875.59
G1	High_{mmp}	3287.70	High_{qspr}	2860.56
	Low_{mmp}	2876.58	Low_{qspr}	3199.60
G2 Constrained	High_{mmp}	5697.19	High_{qspr}	5344.65
	Low_{mmp}	5401.25	Low_{qspr}	5659.71
G2 Unconstrained	High_{mmp}	799.08	High_{qspr}	740.87
	Low_{mmp}	915.07	Low_{qspr}	912.98
G2	High_{mmp}	3153.31	High_{qspr}	3026.46
	Low_{mmp}	4035.52	Low_{qspr}	4113.92
G3 Constrained	High_{mmp}	8819.17	High_{qspr}	8755.27
	Low_{mmp}	7653.87	Low_{qspr}	7656.41
G3 Unconstrained	High_{mmp}	1205.78	High_{qspr}	1226.55
	Low_{mmp}	1253.87	Low_{qspr}	1214.73
G3	High_{mmp}	3736.35	High_{qspr}	3724.87
	Low_{mmp}	3426.48	Low_{qspr}	3417.88

B. Internet Appendix

Table IA.1. Quoted Spreads, True Spreads and Market-Maker Profits
This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0+\beta_y.Year+\beta_1.G_1+\beta_2.G_2+\beta_3.G_3+\beta_4.Event+\beta_5.G_1\times Event+\beta_6.G_2\times Event+\beta_7.G_3\times Event+\beta_8^T.X_{it}+\epsilon_{it}$,
where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.360	12.748	0.308	8.698	0.052	15.712	0.004	8.506
Year	0.013	7.880	0.011	6.540	0.002	6.220	0.000	-1.945
Event	0.003	1.718	0.002	2.182	0.001	1.645	0.000	-0.529
G1	-0.001	-0.229	-0.001	-0.221	-0.001	-1.959	0.000	1.169
G2	0.000	-0.136	0.000	-0.099	0.000	-0.040	0.000	0.741
G3	-0.001	-0.521	-0.001	-0.292	-0.001	-1.927	0.000	-0.120
G1 × Event	0.025	17.716	-0.012	-7.456	0.036	120.890	0.014	72.503
G2× Event	0.021	9.212	-0.017	-9.522	0.038	155.858	0.016	89.870
G3 × Event	0.026	13.311	-0.016	-8.267	0.042	110.453	0.019	77.874
VIX	0.001	5.745	0.001	6.661	0.000	6.151	0.000	-5.749
Turnover	0.000	-1.181	0.000	-1.287	0.000	-1.457	0.000	-0.010
Price	0.068	27.873	0.058	13.471	0.010	29.789	0.000	6.617
Size	-0.039	-15.808	-0.034	-9.855	-0.005	-17.583	0.000	0.585

Table IA.2. Quoted Spreads, True Spreads and Market-Maker Profits

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Constrained Stocks Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.296	6.388	0.251	11.329	0.045	21.746	0.004	6.093
Year	0.012	5.812	0.010	4.433	0.002	4.141	0.000	-6.417
Event	0.005	3.609	0.004	2.289	0.001	3.484	0.000	-1.763
G1	-0.013	-6.700	-0.011	-4.936	-0.002	-4.720	0.000	2.249
G2	-0.016	-5.859	-0.014	-6.649	-0.003	-7.256	0.000	3.118
G3	-0.015	-10.936	-0.013	-6.770	-0.002	-11.924	0.000	4.682
G1 × Event	0.034	27.040	-0.004	-2.064	0.038	190.935	0.014	78.052
G2× Event	0.034	36.980	-0.005	-4.199	0.039	147.316	0.016	77.018
G3 × Event	0.035	37.513	-0.007	-4.937	0.042	155.173	0.018	86.848
VIX	0.001	4.207	0.000	1.273	0.000	2.650	0.000	-0.598
Turnover	0.000	-0.285	0.000	-0.358	0.000	-3.294	0.000	-0.005
Price	0.055	9.680	0.047	18.487	0.009	26.437	0.000	4.792
Size	-0.03	-6.82	-0.027	-12.807	-0.004	-22.208	2.5E-05	4E-01

Table IA.3. Quoted Spreads, True Spreads and Market-Maker Profits

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Unconstrained Stocks Stocks

	Quoted spread		True spread		Market-maker profits		Realized Profits	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.479	15.232	0.420	13.553	0.059	35.624	0.002	2.696
Year	0.015	6.443	0.013	4.552	0.002	7.587	0.000	-2.717
Event	0.008	2.709	0.006	3.115	0.001	3.896	0.000	-1.727
G1	0.012	2.230	0.010	2.358	0.002	4.602	0.000	-0.838
G2	0.013	1.524	0.010	2.262	0.003	6.977	0.000	-0.309
G3	0.013	1.840	0.012	1.584	0.002	2.197	0.000	-4.529
G1 × Event	0.003	0.877	-0.030	-10.715	0.033	71.751	0.014	44.063
G2× Event	0.003	0.635	-0.034	-8.305	0.037	65.062	0.018	80.997
G3 × Event	0.007	2.052	-0.036	-7.817	0.043	53.080	0.022	50.739
VIX	0.001	2.768	0.001	2.273	0.000	4.908	0.000	-4.094
Turnover	0.000	-1.929	0.000	-0.673	0.000	-2.602	0.000	0.081
Price	0.083	13.791	0.072	14.531	0.011	27.460	0.000	1.750
Size	-0.051	-14.798	-0.045	-13.391	-0.006	-27.641	0.000	3.367

Table IA.4. Components of Bid-Ask Spread

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Components with the rounding adjustment: All Stocks					Components without the rounding adjustment: All Stocks				
	Adverse selection		Inventory costs			Adverse selection		Inventory costs	
Coefficient	Estimate	t-stat	Estimate	t-stat	Coefficient	Estimate	t-stat	Estimate	t-stat
Intercept	0.051	5.812	0.257	10.473	Intercept	0.128	10.328	0.232	10.916
Year	0.002	6.819	0.009	6.703	Year	0.005	6.855	0.008	4.714
Event	0.000	-0.491	0.003	2.697	Event	0.000	-0.225	0.003	2.107
G1	0.000	-0.725	0.000	-0.024	G1	0.000	0.170	-0.001	-0.485
G2	0.000	-1.354	0.000	0.192	G2	0.000	0.084	0.000	-0.110
G3	0.000	-0.287	-0.001	-0.255	G3	0.000	0.076	-0.001	-0.527
G1 × Event	-0.002	-12.572	-0.009	-13.047	G1 × Event	0.005	6.094	0.019	16.246
G2× Event	-0.002	-5.609	-0.014	-7.635	G2× Event	0.005	8.352	0.016	11.643
G3 × Event	-0.002	-5.606	-0.014	-11.843	G3 × Event	0.004	5.975	0.021	19.474
VIX	0.000	5.204	0.001	5.649	VIX	0.000	5.512	0.001	4.522
Turnover	0.000	-0.754	0.000	-0.872	Turnover	0.000	-2.539	0.000	-0.591
Price	0.010	9.200	0.047	18.992	Price	0.023	21.984	0.045	17.418
Size	-0.006	-6.552	-0.028	-12.424	Size	-0.014	-12.503	-0.026	-12.450

Table IA.5. Components of Bid-Ask Spread

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Components with the rounding adjustment: Constrained Stocks					Components without the rounding adjustment: Constrained stocks				
	Adverse selection		Inventory costs			Adverse selection		Inventory costs	
Coefficient	Estimate	t-stat	Estimate	t-stat	Coefficient	Estimate	t-stat	Estimate	t-stat
Intercept	0.042	5.992	0.209	8.078	Intercept	0.107	8.692	0.189	9.523
Year	0.002	4.255	0.008	4.815	Year	0.004	6.850	0.008	5.850
Event	0.000	0.769	0.004	3.394	Event	0.001	1.688	0.005	3.225
G1	-0.002	-3.089	-0.009	-3.704	G1	-0.004	-3.449	-0.010	-4.373
G2	-0.003	-14.095	-0.011	-6.378	G2	-0.005	-8.655	-0.012	-7.578
G3	-0.002	-5.943	-0.010	-5.421	G3	-0.005	-8.501	-0.011	-7.322
G1 × Event	-0.001	-2.901	-0.003	-2.054	G1 × Event	0.008	11.993	0.026	16.936
G2× Event	-0.001	-2.005	-0.005	-4.462	G2× Event	0.008	44.559	0.026	39.349
G3 × Event	-0.001	-3.782	-0.006	-5.985	G3 × Event	0.007	31.473	0.028	76.495
VIX	0.000	2.240	0.000	1.894	VIX	0.000	1.568	0.000	2.766
Turnover	0.000	-0.609	0.000	-0.453	Turnover	0.000	-0.615	0.000	-0.512
Price	0.008	10.305	0.038	12.591	Price	0.018	9.837	0.037	12.223
Size	0.00	-6.74	-0.022	-9.043	Size	-0.01	-8.24	-0.020	-9.529

Table IA.6. Components of Bid-Ask Spread

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$,

where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Components with the rounding adjustment: Unconstrained stocks				
	Adverse selection		Inventory costs	
Coefficient	Estimate	t-stat	Estimate	t-stat
Intercept	0.068	8.697	0.352	11.663
Year	0.002	4.052	0.011	4.897
Event	0.001	1.431	0.006	3.200
G1	0.001	1.075	0.008	1.905
G2	0.001	1.345	0.009	1.844
G3	0.002	1.733	0.009	1.905
G1 × Event	-0.006	-9.720	-0.024	-10.182
G2× Event	-0.005	-5.141	-0.028	-9.152
G3 × Event	-0.004	-6.184	-0.032	-10.021
VIX	0.000	2.247	0.001	3.509
Turnover	0.000	-1.462	0.000	-1.283
Price	0.013	11.447	0.059	15.478
Size	-0.008	-9.050	-0.038	-12.714

Components without the rounding adjustment: Unconstrained stocks				
	Adverse selection		Inventory costs	
Coefficient	Estimate	t-stat	Estimate	t-stat
Intercept	0.169	15.148	0.311	16.327
Year	0.006	4.101	0.009	3.305
Event	0.002	2.462	0.006	4.175
G1	0.004	2.256	0.008	2.050
G2	0.003	2.191	0.010	2.576
G3	0.005	1.667	0.009	1.763
G1 × Event	0.000	-0.410	0.003	1.287
G2× Event	0.001	1.027	0.001	0.543
G3 × Event	-0.001	-0.863	0.008	2.032
VIX	0.000	1.624	0.001	2.771
Turnover	0.000	-0.729	0.000	-1.210
Price	0.028	14.148	0.055	18.733
Size	-0.018	-14.271	-0.033	-16.781

Table IA.7. Effective Spreads with Fundamental Prices, mid-Quotes, and Weighted Mid-quotes
This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0+\beta_y.Year+\beta_1.G_1+\beta_2.G_2+\beta_3.G_3+\beta_4.Event+\beta_5.G_1\times Event+\beta_6.G_2\times Event+\beta_7.G_3\times Event+\beta_8^T.X_{it}+\epsilon_{it}$,
where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.279	13.164	0.292	8.072	0.264	16.033
Year	0.008	7.780	0.008	5.544	0.008	5.180
Event	0.002	1.924	0.002	1.647	0.003	2.575
G1	0.000	0.038	0.000	-0.044	0.000	0.009
G2	0.000	-0.090	0.000	-0.162	0.000	0.078
G3	0.000	0.084	0.000	-0.024	-0.001	-0.225
G1 × Event	0.017	25.788	0.015	9.616	0.018	17.831
G2× Event	0.015	10.562	0.013	11.628	0.019	11.193
G3 × Event	0.015	11.401	0.013	8.983	0.024	14.923
VIX	0.001	5.561	0.001	4.959	0.001	4.696
Turnover	0.000	-0.869	0.000	-1.765	0.000	-0.428
Price	0.049	22.700	0.050	11.885	0.048	22.678
Size	-0.030	-15.206	-0.031	-8.763	-0.028	-18.358

Table IA.8. Effective Spreads with Fundamental Prices, mid-Quotes, and Weighted Mid-quotes
This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0+\beta_y.Year+\beta_1.G_1+\beta_2.G_2+\beta_3.G_3+\beta_4.Event+\beta_5.G_1\times Event+\beta_6.G_2\times Event+\beta_7.G_3\times Event+\beta_8^T.X_{it}+\epsilon_{it}$,
where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Constrained Stocks

	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.228	13.243	0.239	11.230	0.216	8.728
Year	0.007	5.608	0.007	4.029	0.007	4.972
Event	0.003	4.030	0.003	3.510	0.004	2.420
G1	-0.008	-2.413	-0.009	-3.619	-0.008	-2.883
G2	-0.011	-9.266	-0.011	-4.742	-0.011	-8.945
G3	-0.010	-6.942	-0.010	-7.881	-0.010	-14.706
G1 × Event	0.025	22.199	0.021	16.907	0.025	12.993
G2× Event	0.025	76.437	0.021	38.504	0.027	15.938
G3 × Event	0.023	117.106	0.020	92.987	0.030	31.777
VIX	0.000	2.061	0.000	1.831	0.000	1.557
Turnover	0.000	-0.527	0.000	-0.622	0.000	-0.561
Price	0.04	15.07	0.040	15.633	0.039	11.435
Size	-0.02	-13.29	-0.024	-11.891	-0.022	-8.914

Table IA.9. Effective Spreads with Fundamental Prices, mid-Quotes, and Weighted Mid-quotes
This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$,
where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Unconstrained Stocks

	Mid-quote effective spreads		Hagstormer effective spreads		True price effective spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.382	13.830	0.388	14.474	0.355	14.209
Year	0.010	3.849	0.010	3.855	0.010	4.529
Event	0.005	2.930	0.005	3.185	0.006	3.416
G1	0.007	1.323	0.008	2.623	0.008	2.486
G2	0.007	1.886	0.008	1.076	0.009	1.568
G3	0.010	1.811	0.010	1.418	0.009	2.055
G1 × Event	0.000	0.049	0.001	0.362	0.004	1.257
G2× Event	0.001	0.204	0.001	0.237	0.006	1.997
G3 × Event	-0.002	-0.650	-0.001	-0.391	0.013	4.308
VIX	0.001	2.544	0.001	2.062	0.001	2.056
Turnover	0.000	-0.683	0.000	-1.017	0.000	-1.242
Price	0.062	13.496	0.062	15.741	0.059	13.477
Size	-0.040	-13.514	-0.040	-15.020	-0.037	-13.363

Table IA.10. Realized spreads

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$,
where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Mid-quote realized spreads		True price realized spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat
Intercept	0.101	7.324	0.098	8.286
Year	0.000	0.259	0.000	0.581
Event	0.000	-0.062	0.000	-0.085
G1	0.001	0.561	0.001	0.925
G2	0.001	0.468	0.001	0.741
G3	-0.001	-2.570	-0.001	-2.877
G1 × Event	0.007	5.236	0.008	5.025
G2× Event	0.008	8.885	0.010	6.765
G3 × Event	0.005	8.193	0.011	42.925
VIX	0.000	-1.262	0.000	-2.059
Turnover	0.000	-0.403	0.000	-1.160
Price	0.010	6.916	0.010	8.127
Size	-0.009	-6.220	-0.008	-6.998

Table IA.11. Realized spreads

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0+\beta_y.Year+\beta_1.G_1+\beta_2.G_2+\beta_3.G_3+\beta_4.Event+\beta_5.G_1\times Event+\beta_6.G_2\times Event+\beta_7.G_3\times Event+\beta_8^T.X_{it}+\epsilon_{it}$, where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Constrained Stocks

	Mid-quote realized spreads		True price realized spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat
Intercept	0.072	4.189	0.069	4.740
Year	0.000	0.344	0.000	-0.339
Event	0.000	-0.168	0.000	-0.190
G1	-0.004	-2.662	-0.003	-2.720
G2	-0.005	-6.498	-0.004	-4.146
G3	-0.005	-7.710	-0.005	-3.687
G1 × Event	0.010	7.560	0.010	5.590
G2× Event	0.011	56.598	0.011	5.589
G3 × Event	0.009	5.717	0.012	6.226
VIX	0.000	-1.257	0.000	-2.714
Turnover	0.000	-1.026	0.000	-1.669
Price	0.007	3.550	0.006	4.198
Size	-0.01	-3.26	-0.005	-3.790

Table IA.12. Realized spreads

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$,
where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Unconstrained Stocks

	Mid-quote realized spreads		True price realized spreads	
Coefficient	Estimate	t-stat	Estimate	t-stat
Intercept	0.133	11.145	0.122	6.204
Year	0.000	0.023	0.000	-0.216
Event	-0.001	-0.383	-0.001	-0.396
G1	0.013	3.555	0.013	2.583
G2	0.012	1.637	0.011	1.703
G3	0.009	6.399	0.007	2.513
G1 × Event	0.000	-0.050	0.002	0.446
G2× Event	0.005	1.030	0.009	3.420
G3 × Event	0.001	0.807	0.010	2.279
VIX	0.000	-1.442	0.000	-2.355
Turnover	0.000	-0.595	0.000	-0.389
Price	0.013	7.720	0.012	5.597
Size	-0.011	-8.241	-0.010	-5.607

Table IA.13. Proportions of Price Discovery

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$,
where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Discovery through market orders		Discovery through limit orders		Discovery through new info	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.204	10.840	0.342	16.241	0.546	215.660
Year	0.009	4.529	-0.008	-5.047	0.001	2.321
Event	0.003	1.531	-0.003	-1.313	0.000	-1.016
G1	-0.007	-1.631	0.006	2.295	-0.001	-2.409
G2	-0.002	-0.507	0.001	0.418	0.000	-0.884
G3	-0.003	-1.364	0.002	0.715	0.000	-0.306
G1 × Event	-0.043	-8.399	0.033	6.112	-0.010	-16.292
G2× Event	-0.047	-14.586	0.038	16.095	-0.008	-12.975
G3 × Event	-0.040	-14.868	0.032	9.695	-0.008	-14.860
VIX	-0.001	-2.360	0.001	3.177	0.000	0.524
Turnover	0.000	0.360	0.000	-0.303	0.000	-0.344
Price	0.026	23.388	-0.016	-15.525	0.009	42.194
Size	-0.014	-8.930	0.010	6.258	-0.004	-21.456

Table IA.14. Proportions of Price Discovery

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Constrained Stocks

	Discovery through market orders		Discovery through limit orders		Discovery through new info	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.212	7.820	0.335	22.054	0.548	140.716
Year	0.011	4.626	-0.010	-4.663	0.002	2.489
Event	0.006	2.125	-0.006	-2.896	0.000	0.055
G1	-0.007	-1.546	0.006	1.771	-0.001	-2.693
G2	-0.004	-2.402	0.004	1.003	-0.001	-0.780
G3	-0.001	-0.341	0.001	0.416	0.000	0.062
G1 × Event	-0.043	-5.998	0.032	5.293	-0.011	-17.047
G2× Event	-0.051	-16.103	0.041	20.097	-0.010	-10.664
G3 × Event	-0.046	-9.499	0.037	9.846	-0.010	-15.409
VIX	-0.001	-3.178	0.001	2.970	0.000	-0.206
Turnover	0.000	0.361	0.000	-0.183	0.000	-0.305
Price	0.026	11.382	-0.017	-13.162	0.009	22.736
Size	-0.015	-6.453	0.010	8.132	-0.004	-13.843

Table IA.15. Proportions of Price Discovery

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$,

where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Unconstrained Stocks

	Discovery through market orders		Discovery through limit orders		Discovery through new info	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	0.176	6.874	0.364	13.626	0.540	115.918
Year	0.009	3.658	-0.008	-2.913	0.001	2.783
Event	0.004	1.481	-0.004	-1.484	0.000	-0.155
G1	-0.004	-1.289	0.004	1.190	0.000	-0.481
G2	0.003	0.840	-0.003	-0.821	0.000	0.522
G3	-0.004	-0.731	0.004	0.907	0.000	-0.294
G1 × Event	-0.044	-16.124	0.036	16.940	-0.008	-10.842
G2× Event	-0.041	-8.459	0.034	8.987	-0.006	-9.707
G3 × Event	-0.027	-6.149	0.022	5.177	-0.005	-4.646
VIX	-0.001	-3.117	0.001	3.140	0.000	-0.556
Turnover	0.000	0.238	0.000	-0.363	0.000	-0.315
Price	0.023	11.957	-0.015	-5.636	0.008	27.772
Size	-0.011	-5.020	0.007	3.404	-0.003	-9.933

Table IA.16. Speed of Price Discovery

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:
$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

All Stocks

	Delay Measure D1		Delay Measure D2		Delay Measure D3	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	2.523	27.042	1.688	27.301	1.677	33.192
Year	0.029	1.363	0.010	1.130	0.010	1.093
Event	0.079	3.416	0.036	3.213	0.039	3.333
G1	-0.011	-0.705	-0.006	-0.847	-0.006	-0.748
G2	0.007	0.444	0.003	0.568	0.003	0.351
G3	0.010	0.656	0.003	1.002	0.003	0.444
G1 × Event	0.087	8.294	0.037	4.272	0.037	4.185
G2× Event	0.071	5.740	0.029	4.455	0.029	3.571
G3 × Event	0.092	7.524	0.037	8.380	0.037	5.981
VIX	-0.014	-3.351	-0.007	-5.925	-0.007	-4.093
Turnover	0.000	-0.615	0.000	-0.462	0.000	-0.612
Price	0.027	2.469	0.015	3.173	0.014	4.087
Size	-0.143	-17.529	-0.062	-11.549	-0.061	-15.100

Table IA.17. Speed of Price Discovery

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0 + \beta_y.Year + \beta_1.G_1 + \beta_2.G_2 + \beta_3.G_3 + \beta_4.Event + \beta_5.G_1 \times Event + \beta_6.G_2 \times Event + \beta_7.G_3 \times Event + \beta_8^T.X_{it} + \epsilon_{it}$, where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Constrained Stocks

	Delay Measure D1		Delay Measure D2		Delay Measure D3	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	2.418	19.534	1.652	26.264	1.641	41.306
Year	0.017	0.756	0.006	0.609	0.006	0.720
Event	0.080	4.823	0.036	3.989	0.039	4.776
G1	-0.052	-2.806	-0.026	-3.032	-0.026	-3.468
G2	-0.038	-1.927	-0.016	-2.432	-0.016	-2.292
G3	-0.030	-2.277	-0.015	-1.990	-0.015	-2.127
G1 × Event	0.129	5.913	0.058	5.545	0.058	5.698
G2× Event	0.122	7.139	0.049	7.648	0.050	8.088
G3 × Event	0.133	8.783	0.057	7.997	0.057	8.822
VIX	-0.015	-3.538	-0.007	-4.926	-0.007	-6.774
Turnover	0.000	-0.817	0.000	-0.971	0.000	-0.849
Price	0.008	0.870	0.007	1.425	0.006	1.480
Size	-0.129	-13.239	-0.057	-11.340	-0.056	-18.586

Table IA.18. Speed of Price Discovery

This table presents the difference-in-differences coefficients that quantify the average changes in liquidity measures due to the TSP program. The difference-in-differences coefficients are based on the following panel regression specification at the millisecond TAQ-level for each stock:

$MQ_{it} = \beta_0+\beta_y.Year+\beta_1.G_1+\beta_2.G_2+\beta_3.G_3+\beta_4.Event+\beta_5.G_1\times Event+\beta_6.G_2\times Event+\beta_7.G_3\times Event+\beta_8^T.X_{it}+\epsilon_{it}$,

where MQ_{it} is a market quality measure for stock i at time t (millisecond-level); $Year$ is a dummy that equals one for all trades in the year 2018; G_1, G_2, G_3 are dummies that equal one for stocks belonging to the respective test groups; $Event$ is a dummy that equals one for all trades during the TSP regime; X_{it} is the set of exogenous variables containing the VIX index, stock turnover, stock price, and size that are available at beginning of trade t's transaction day. Standard errors are clustered by firm and day. Constrained (unconstrained) stocks contain only a subset of stocks whose average quoted bid-ask spreads are lower (higher) than 5 cents during the non-TSP regime. Our TSP regime data spans from November 1, 2016 to November 30, 2016 and from September 1, 2018 to September 30, 2018. The non-TSP regime data spans from September 1 2016 to September 30 2016 and from November 1, 2018 to November 30, 2018.

Unconstrained Stocks

	Delay Measure D1		Delay Measure D2		Delay Measure D3	
Coefficient	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	2.672	34.417	1.752	29.277	1.741	44.851
Year	0.003	0.151	0.002	0.250	0.002	0.186
Event	0.072	3.372	0.035	4.197	0.037	3.661
G1	0.075	5.917	0.036	4.970	0.036	5.031
G2	0.062	4.099	0.026	2.665	0.026	2.966
G3	0.077	5.812	0.037	4.219	0.037	3.985
G1 × Event	-0.019	-1.220	-0.014	-1.892	-0.015	-1.726
G2× Event	-0.019	-1.577	-0.007	-1.013	-0.008	-1.249
G3 × Event	-0.001	-0.075	-0.009	-1.258	-0.009	-1.040
VIX	-0.014	-4.649	-0.007	-4.428	-0.007	-5.667
Turnover	0.000	-0.467	0.000	-0.559	0.000	-0.712
Price	0.047	5.114	0.022	4.749	0.021	5.129
Size	-0.158	-23.442	-0.068	-14.881	-0.067	-20.994