THE LIQUIDITY-MAXIMIZING PRICE OF A STOCK: A TALE OF TWO DISCRETENESSES*

SIDA LI AND MAO YE

Economists commonly assume that price and quantity are continuous variables, while in reality both are discrete. As U.S. regulations mandate a one-cent minimum tick size and a 100-share minimum lot size, we predict that less volatile and more active stocks will choose higher prices to make pricing more continuous and quantity more discrete. Despite heterogeneous optimal prices, all firms achieve their liquidity-maximizing prices when their bid–ask spreads equal two ticks, i.e. when frictions from discrete pricing equal those from discrete lots. Empirically, our theoretical model explains 57% of cross-sectional variations in stock prices and 81% of cross-sectional variations in bid–ask spreads. The adjustment toward liquidity-maximizing prices rationalizes 91% of stock splits and contribute 94 bps to split announcement returns. Liquidity-maximizing pricing could increase median U.S. stock value by 106 bps and total U.S. market capitalization by $93.7 billion. *JEL Codes: G10, G14, G18, D47*

* We thank James Angel (discussant), Kerry Back, Malcolm Baker, Markus Baldauf (discussant), Dan Bernhardt, Hendrik Bessembinder, Kevin Crotty, John Campbell, Charles Cao, Xavier Gabaix, Robin Greenwood, Albert (Pete) Kyle, Giang Nguyen, Maureen O'Hara, Shiyang Huang (discussant), Gideon Saar, Andrei Shleifer, Chester Spatt, Jeremy Stein, Bart Zhou Yueshen, Xingtan Zhang, and seminar participants at Harvard University, Rice University, the University of Illinois, Penn State University, Cornell University, the Market Microstructure Exchange, EFA 2021, SAFE 2021, the Econometric Society North America Summer Meeting 2021, Econometric Society Australasia 2021, the 11th Annual 2022 Stern/Salomon Center Microstructure Conference, CICF 2022, and the UT Austin PhD Student Symposium for their helpful comments. This research is supported by National Science Foundation grant 1838183. Computational power is provided by the Extreme Science and Engineering Discovery Environment (XSEDE) project SES120001.
Corresponding author, Mao Ye: Cornell University and NBER, 365 Sage Hall, 114 E Ave, Ithaca, NY, 14853; phone: 607-255-9509; email: my87@cornell.edu.

1. INTRODUCTION

Price and quantity are two of the most important variables in economics. In most economic models, price and quantity are continuous variables, but they are discrete in reality, even in most liquid markets such as U.S. stock exchanges. Regulation National Market System (Reg NMS) mandates a minimum price variation (the tick size) of one cent for stocks priced above $1 per share. Reg NMS also defines the minimum quantity of shares needed to establish a bid or offer as one round lot, which for most stocks is 100 shares.[1] A U.S. firm can, therefore, choose a high price per share for a more continuous price but a more discrete quantity or a low price per share for a more discrete price but a more continuous quantity. This paper shows that the tradeoff between these two seemingly small economic frictions (discrete price and discrete quantity) rationalizes several well-known puzzles in market microstructure, behavioral finance, and corporate finance.

We unlock these puzzles in three steps. First, we show that a stock's nominal share price significantly impacts stock liquidity when price and quantity are discrete. This result explains why stocks with similar fundamentals can have dramatically different liquidity and why stock splits can significantly impact liquidity. A parsimonious three-variable model generated by our theoretical predictions can explain more than 80% of cross-sectional variation in stock liquidity. Second, the tradeoff between discrete price and discrete quantity implies that each firm has an optimal price per share. We provide the analytical solution for the liquidity-maximizing price, which explains 57% of cross-sectional variations in stock prices. Third, we show that adjustment towards the liquidity-maximizing price explains more than 90% of stock splits. The improved liquidity then increases corporate value by 94 bps, which is more than one-third of the split-announcement return of 273 bps.

[1] See the final rule under Reg NMS issued by the U.S. Securities and Exchange Commission (SEC) Release No. 34-51808. Reg NMS offers some exemptions such that brokers may internalize their customers' order flows at sub-penny prices and customers can trade fractions of shares on some occasions. The bid and ask price of a stock is, however, bounded by tick size and lot sizes. Also, these exemptions do not change the economic trade-offs modeled by our paper, as the tick (lot) constraint remains higher for low- (high-) priced stocks.

We first set up a model that quantifies the total frictions generated by the discrete price and the discrete quantity. We find that the lot size constraint is proportional to the nominal price. Our model first explains why most stocks in our sample have a median depth of one lot. If new information arrives and the market maker is adversely selected by better-informed traders, her loss doubles if she quoted two instead of one lot. The binding one lot implies that the market maker's loss increases linearly with the share price. To compensate for the loss from the lot constraint, the market maker sets a lot-driven percentage spread that increases linearly with the share price p. On the other hand, we find that the tick size constraint is inversely proportional to p. Intuitively, discrete pricing widens the competitive bid–ask spread to the next available tick grid. We find that the widening effect equals one tick as long as the stock value follows a smooth and continuous distribution. The uniform friction of one tick for stocks at all prices implies that the tick size constraint increases linearly with $\frac{1}{p}$.

Insofar as the total frictions generated by discreteness equal the sum of a function linear in p and a function linear in $\frac{1}{p}$, either a price that is too low or too high harms liquidity. We show that differences in share prices can quantitively explain why stocks with similar fundamentals can have dramatically different levels of liquidity. For example, the very low price of Ford (\$7) explains why its percentage spread (14 bps) is four times higher than that of GM (priced at \$30). The very high price of Amazon (\$3,305) explains why its percentage spread (4.62 bps) was six times higher than that of Microsoft (priced at \$255).

More generally, our model explains more than 80% of cross-sectional variation in bid–ask spreads using only three variables: volatility, dollar trading volume, and nominal price. Surprisingly, the R^2 of our parsimonious model is significantly higher than in existing benchmarks (Madhavan 2000; Stoll 2000), even though we use only a subset of their explanatory variables. Like Madhavan's (2000) and Stoll's (2000) models, our model includes volatility and dollar trading volume. Therefore, our increase in R^2 comes from the functional form of p. Madhavan (2000) controls for $\frac{1}{p}$ while Stoll (2000) controls for $log(p)$. Both specifications impose a monotonic relationship between the nominal price and liquidity, but we find that their true relationship is U-shaped. We find that the R^2 in Madhavan (2000) would rise from 0.62 to 0.81 and the R^2 in Stoll (2001) would rise from 0.65 to 0.82 if they adopted the functional form implied by our model: subtracting one tick from the bid–ask spread to control for the tick-size constraint and then using the log of the price as an independent variable to control for the lot-size constraint.

In our second step, we allow firms to choose an optimal price that minimizes the total frictions from tick- and lot-size constraints. Our main theoretical result in this step is the Two-Tick Rule: all firms maximize their liquidity when their bid–ask spreads are two ticks wide. Intuitively, as the friction caused by a discrete price equals one tick for stocks at any price, all firms reach their liquidity-maximizing prices (we use "optimal prices" for short hereafter) when the friction generated by the discrete lot also equals one tick, i.e., when their bid–ask spreads equal two ticks.

Empirically, we find that the Two-Tick Rule can explain 57% of cross-sectional variation in share prices using only two variables: volatility and dollar volume. The homogenous two-tick optimal bid–ask spread implies heterogeneous share prices. For example, our model predicts that stocks with higher volatility will choose a lower share price. The first way to understand this result is through the percentage spread. As an increase in volatility increases the percentage spread, stocks with higher volatility achieve their two-tick optimal spreads at lower prices. The second way to explain this result is through lot constraints. An increase in volatility increases adverse selection risk for market makers, making firms reduce their share prices to reduce the adverse-selection risk their market makers face. The third way to explain the negative correlation between volatility and price is through tick constraints. An increase in volatility implies that price movement is constrained to a lesser extent by the tick size, which gives firms more room to choose lower prices. The same three intuitions also explain why higher-dollar-volume stocks should choose higher prices. An increase in dollar volume increases market makers' revenue and reduces the percentage spread. Therefore, higher-volume stocks reach their two-tick optimal spreads at higher prices. An alternative way to understand this result is that an increase in volume makes tick constraints more salient than lot constraints, giving firms the incentive to increase their share prices.

In addition to its great explanatory power, the Two-Tick Rule also rationalizes two puzzles in the behavioral finance literature. Baker, Greenwood, and Wurgler (2009) find it puzzling that volatile firms are more likely to split their stocks because they have a *"greater chance of reaching a low price anyway."* Shue and Townsend (2018) provide a behavior-based interpretation. As investors think partly about stock-price changes in dollars rather than percentage units, low-priced stocks should experience more extreme return responses to news. Shue and Townsend (2018) show that lower share prices can lead to higher volatility, and we show that stocks with higher volatility choose lower prices. Weld et al. (2009) find that firms choose prices that are similar to those chosen by similarly sized firms and industry peers, and they explain this puzzle using social norms. The Two-Tick Rule provides an alternative interpretation of the clustering. We show that firms that operate in the same industry may choose similar prices because their volatilities are similar. Our model also rationalizes price clustering around size, as an increase in size or turnover increases dollar volume. An increase in dollar volume reduces the percentage spread and makes the tick size a more binding constraint. Therefore, in addition to rationalizing price clustering, our paper also explains why large stocks cluster at higher prices.

As firms can adjust their nominal prices using stock splits, the last step in our analysis addresses three questions in the corporate finance literature. 1) Why do firms split their stocks? 2) What explains the positive returns that follow splits? 3) Why are these returns realized mostly on announcement dates but not ex dates (Fama et al. 1969; Grinblatt, Masulis, Titman 1984)?[2] Our model explains more than 91% of splits: 1,077 of 1,196 stock splits adjust nominal price towards the optimal price implied by the Two-Tick Rule. Among the 107 "incorrect" splits, 74 make the correct decision to split, except that they choose an overly aggressive split ratio.

Our tick-and-lot channel helps us rationalize the puzzles raised by two canonical channels for stock splits: the signaling channel and the trading range channel. In the

[2] Fama et al. (1969) do not find any return effects during ex dates. Grinblatt, Masulis, Titman (1984) find return effects on ex dates using more granular data, but the return effects on ex dates is only one-third the size of the effects on announcement dates.

signaling channel (Brennan and Copeland, 1988), managers use splits to signal good news. Fama et al. (1969), Lakonishok and Lev (1987), and Asquith, Healy, and Palepu (1989) find, however, that stock prices, earnings, and profits increase significantly before splits but not after splits. Their results are inconsistent with the signaling channel but consistent with our tick-and-lot channel. A previous increase in a stock price increases the lot-size constraint on that stock, and stock splits relieve this constraint. The trading range hypothesis (Baker and Gallagher 1980; Lakonishok and Lev 1987) maintains that managers split stocks to achieve optimal trading ranges. An open question is what decides the optimal trading range and why larger stocks have higher optimal trading ranges (Stoll and Whaley 1983). Our Two-Tick Rule provides the analytical solution to the trading range, and the uniform two-tick bid–ask spread implies that larger stocks should choose higher prices because they have smaller percentage spreads.

Fama et al. (1969) find that market judgments concerning splits are fully reflected in prices almost immediately after announcement dates. They regard this result as evidence that the stock market is efficient because prices adjust rapidly to new information. What this new information contained in split announcements consists in is an open question. Mcnichols and Dravid (1990) find that the explanatory power of nominal prices is considerably greater than earning forecast errors, suggesting that nominal prices are more fundamental to split factor choices than managers' private information. We address these questions by showing that the new information comes from the announced split ratio, as we can use this ratio to predict a stock's liquidity change along with its current price and bid–ask spread. We find that the predicted liquidity changes on announcement dates match almost one to one with the realized liquidity changes on ex dates. As most stock splits move nominal prices towards the optimum, the predicted percentage spread drops by 15.22 bps. As an increase in liquidity increases corporate value, a one bps reduction in the predicted percentage spread increases that value by 6.18 bps. Therefore, correct split ratios increase corporate value by 94 bps, which explains more than one-third of the split-announcement return of 273 bps.

We base our main analysis on uniform tick and lot sizes because of current U.S.

regulations, and in Section 5 we evaluate the policy implications of alternative tick and lot systems. As the uniform system imposes the same tick and lot sizes for all stocks, one possibility is to move to a proportional system such that lower-priced stocks have smaller tick sizes and larger lot sizes. Surprisingly, we show that a proportional system would reduce liquidity. The intuition is as follows. The uniform system seems like a "one-size-fits-all" system, but it allows a firm to choose optimal prices to balance discrete prices and quantities. The proportional system harms liquidity because it destroys this degree of freedom. Therefore, the proportional system is the true "one-size-fits-all" system because it imposes a uniform level of discreteness on stocks with heterogeneous characteristics.

As the first study to examine discreteness in both pricing and quantity, our paper offers significantly different predictions from those implied by models where only one variable is continuous. The closest paper to ours is Budish, Cramton, and Shim (2015, "BCS" hereafter), who consider a market with discrete quantities but continuous pricing. BCS find that speed races can lead to a positive percentage spread even if all information is public because a high-frequency trader can snipe stale quotes if he is the first to react to the public information. We find that lot-size constraints are one driver of this positive percentage spread: the percentage spread in BCS converges to zero if the dollar lot size converges to 0. Therefore, our paper shows two ways to reduce the percentage spread under the BCS framework. The first way is to allow regulators to reduce the lot size. The second way is for firms to reduce their share prices through stock splits. In fact, firms can reduce their bid–ask spreads to zero under the BCS framework if they split their shares towards zero per share. In our model, the economic factor that prevents such aggressive splits is discrete pricing. Our model predicts that the recent policy initiative to reduce tick and lot sizes (Gensler 2022) will improve liquidity.

Angel's (1997) optimal tick-size hypothesis focuses only on price discreteness. In his framework, firms can perfectly neutralize a twofold increase in the tick size through a 2-for-1 reverse split. By adding discrete quantities, we find that a 2-for-1 reverse split does not neutralize a twofold increase in the tick size. More surprisingly, a 2-for-1 reverse split leads to the same outcome as doing nothing at all. We find that the total frictions generated by tick and lot sizes equal their product. Although a 2-for-1 reverse split retains the original relative tick size, it also doubles the lot-size constraints, leaving their total frictions and the percentage spread unchanged.

2. MODEL

In this section, we set up a three-stage model, where the regulator, the firm, and traders make decisions sequentially.

Stage 1: The regulator's decision: The regulator moves first and sets the tick size Δ and lot size L at time $t = -2$. In the benchmark model we present in Section 3, we consider continuous prices and discrete lots. Our main analysis in Section 4 reflects the uniform system in the U.S., where all stocks have the same discrete tick and lot sizes. In Section 5, we compare the uniform system with proportional tick and lot sizes, which incorporates regulations enforced in other jurisdictions.

***Stage 2: The firm's decision*:** The firm starts at fundamental value v at time $t = -1$. The firm chooses its price per share p and shares outstanding h such that $p = \frac{v}{h}$.[3] The firm's value v_t then continuously evolves over $t \in (-1, +\infty)$ as a Poisson jump process. The intensity of the jump is $\lambda_J > 0$. The size of the jump is σv_t or $-\sigma v_t$ at equal probability, so v_t is a martingale. The market opens at $t = 0$ and $t \in (-1,0)$ reflects the implementation period. When the firm chooses h through an IPO or stock split, it usually takes about a month to implement the change. Therefore, the firm knows only the distribution of its initial trading price when choosing h. The firm's objective is to maximize its expected liquidity or, equivalently, minimize its traders' expected transaction costs over $t \in (0, +\infty)$.[4]

---

[3] We allow h to be a continuous variable. As Reg NMS allows traders to establish quotes in mixed lots such as lots of 101 shares, the true binding constraints in reality and in our model mean that a quote cannot be smaller than one lot in size.

[4] We limit the scope of our paper to solving the nominal price that maximizes liquidity. In reality, firms have other considerations when determining their nominal prices. For example, managers may maintain a high

Stage 3: The traders' decision. The stock's transaction costs are determined by three types of traders: a competitive market maker, uninformed traders, and informed traders.

Uninformed traders aim to trade an exogenous fraction λ_I of the firm per unit of time, and they aim to minimize their transaction costs by choosing how they slice and dice their parent orders into a series of child orders. We call this choice $\{\lambda_q | q \in N^+\}$, where each λ_q is the Poisson arrival rate of child orders of q lots. They choose all λ_q subject to the constraint that total liquidity demand is $\lambda_I h$ shares per unit time, i.e., $\sum_{q=1}^{\infty} qL\lambda_q = \lambda_I h$. For example, uninformed traders can choose

$$\lambda_q = \begin{cases} \frac{\lambda_I h}{L}, & for\ q = 1 \\ 0, & for\ q \geq 2 \end{cases}. \tag{1}$$

In this case, all uninformed orders arrive with the minimum round lot. We assume symmetric demand for buy and sell liquidity, i.e., the frequency of buy and sell orders of size q are both $\frac{\lambda_q}{2}$.

The market maker chooses the ask and bid prices as $\{A_t^i, B_t^i\}$, where i stands for the price for the i^{th} lot. The dollar transaction cost for a buy order sized q is

$$C_B(q) = \sum_{i=1}^{q} (A_t^i - p_t)L. \tag{2}$$

The dollar transaction cost for a sell order sized q is

$$C_S(q) = \sum_{i=1}^{q} (p_t - B_t^i)L. \tag{3}$$

The uninformed traders minimize the total transaction cost per unit time

price to exclude small investors (Buffett 1983, Schultz 2000), or they may cater to the market by splitting their shares when investors place higher valuations on low-price firms (Baker, Greenwood, and Wurgler 2009). Fortunately, our model can solve for bid–ask spreads given any nominal prices, and firms that deviate from liquidity-optimal prices provide a great laboratory for testing our liquidity model, which we do in Sections 6 and 8. Our back-of-the-envelope calculation in Section 8 estimates the cost of liquidity when firms use nominal prices to pursue objectives other than maximizing liquidity.

$$\min_{\{\lambda_q | q \in N^+\}} \sum_{q=1}^{\infty} [C_B(q) + C_S(q)] \frac{\lambda_q}{2},$$

$$s.t.\ \sum_{q=1}^{\infty} qL\lambda_q = \lambda_I h. \quad (4)$$

Informed traders know the value of the stock before each jump, and they profit from adversely selecting the market maker. Upon arrival, they sweep all outstanding bids above the fundamental value or all asks below the fundamental value. Thus, informed traders' profit is proportional to the outstanding depth. There are two ways to interpret the adverse-selection risk in our model. First, v_t is common knowledge, but the market maker fails to cancel the stale quote. In this case, the market maker in our model is equivalent to the liquidity-providing HFT in BCS, and informed traders are equivalent to the stale-quote-sniping HFTs in BCS. Second, v_t is private information but is revealed after each trade (Baldauf and Mollner, 2020; Admati and Pfleiderer 1988; Anshuman and Kalay 1998). Both scenarios lead to the same model. For the sake of tractability, we assume that informed traders can adversely select the market maker only once per piece of information. Without this simplification, the optimization problem for the firm is not well-defined because the bid–ask spread would be a nonstationary function over time.[5] All other firm fundamentals, σ, λ_I, and λ_J, are public information for traders and the firm.

After observing $\{\lambda_q\}$, the market maker quotes competitive prices on the bid and ask sides of the market. Her choice variable is a set of bid and ask prices $\{A_t^i, B_t^i\}$, where A_t^i and B_t^i can be any number when pricing is continuous, but they need to be integer multiples of Δ when pricing is discrete.

3. CONTINUOUS PRICING AND DISCRETE LOTS

In this section, the regulator sets a tick size of $\Delta = 0$ and a lot size of $L > 0$. We solve the model through backward induction. In subsection 3.1, we solve traders' optimal choices

[5] Glosten and Milgrom (1985), Vayanos (1999), and Back and Baruch (2004) characterize these non-stationary bid–ask spreads, although their solutions are either very complicated or available only numerically.

given share prices and tick and lot sizes. If we enforce $h = 1$ and assume that adverse selection comes from public information, our model degenerates into BCS. Therefore, BCS is a special case of our model where the firm does not optimize its price and the regulator chooses suboptimal tick and lot sizes. In subsection 3.2, we discuss the firm's choice of h.

3.1 Traders' Choice

Proposition 1 shows the optimal strategies for uninformed traders and the market maker. Uninformed traders' optimal strategy is to submit a series of child orders of one lot each. The market maker always displays one lot at the best bid and offer (BBO) and quickly refills another lot when the original lot is consumed. Intuitively, uninformed traders slice to the minimal lot because larger orders execute at worse prices. The proof of Proposition 1 shows this intuition in two steps.

First, uninformed traders would not slice parent orders into child orders with heterogeneous sizes because larger orders walk up and book and execute at worse prices. Suppose uninformed traders choose two sizes, q_1 and q_2 ($q_1 < q_2$).[6] Then the competitive market maker would quote two tiers of liquidity. She offers a better quote for the first q_1 lots because this quote executes against uninformed orders of sizes q_1 and q_2. The competitive market maker then offers a worse quote for the next $(q_2 - q_1)$ lots because it executes only against orders of size q_2. Therefore, q_2 child orders execute at worse prices than q_1 child orders because the former walk up the book. Therefore, a profitable deviation for uninformed traders is to reduce the child order size from q_2 to q_1.

Second, conditional on a homogeneous child order size, only child order sizes of one minimum lot, i.e., $q = 1$, can sustain the equilibrium. In this case, the market maker can quote only one lot at the bid and ask, which minimizes her adverse-selection risk and thereby the bid–ask spread. Otherwise, suppose uninformed traders slice parent orders into child orders of two lots. Then, the competitive market maker needs to maintain a quote for two lots. However, the market maker's quotes will be entirely adversely selected when the price jumps, so the market maker's loss-per-jump is proportional to her displayed quote

[6] If there are more than two order sizes, denote the two smallest order sizes as q_1 and q_2.

size. Thus, an increase in quote size increases her loss during adverse selection, so the market maker's break-even spread widens. An alternative way to understand this result is that a decrease in child order size q_1 increases the order arrival rate λ_{q_1}. The arrival rate of uninformed child orders reaches its maximum ($\frac{\lambda_I h}{L}$) when all child orders are one lot. As the intensity of adverse-selection λ_J is a constant, an increase in the arrival rate of uninformed child orders reduces the break-even bid–ask spread.

Our model predicts trading in minimum lots, which matches the empirical facts. O'Hara, Yao, and Ye (2014) find that more than 50% of trades are sized at exactly 100 shares, and we find that this ratio is as high as 87.5% when the bid–ask spread is not bounded by one tick.

Enjoying the benefit of slicing orders to the minimum lot, however, requires technology that makes it possible to slice the parent order into many child orders. Therefore, our model rationalizes algorithmic traders who are slower than HFTs.[7] Brogaard et al. (2015) document the existence of "SlowColos," who co-locate at a stock exchange but are slower than HFTs. Yet, it is unclear why they choose to be fast but not the fastest. We conjecture that execution algorithms constitute one type of SlowColo: they need to be fast enough to slice many child orders in a short time, but they do not need to be the fastest to select other traders adversely or to avoid being adversely selected by the fastest traders.

The next step in solving the equilibrium is to pin down the transaction cost, measured by the percentage spread $\mathcal{S}_t^L = \frac{s_t^L}{p_t}$. $\mathcal{S}_t^L$ equates the revenue from uninformed traders with the loss from informed traders. The dollar arrival rate for uninformed traders is $\lambda_I v_t \equiv \lambda_I p_t h$, and the revenue required to provide liquidity to uninformed traders is $\frac{\mathcal{S}_t^L}{2}$ basis points. The intensity for an order to be sniped is λ_J and the marker maker loses $p_t L \cdot \left(\sigma - \frac{\mathcal{S}_t}{2}\right)$ when she is sniped. Therefore, the equilibrium percentage spread solves

[7] To the best of our knowledge, the only other interpretation has been offered in a companion paper (Li, Wang, and Ye 2021), which shows that slow traders use execution algorithms to choose between market and limit orders.

$$\lambda_I p_t h \cdot \frac{\mathcal{S}_t^L}{2} = \lambda_J p_t L \cdot (\sigma - \frac{\mathcal{S}_t}{2}). \tag{5}$$

The solution is

$$\mathcal{S}_t^L = \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L}. \tag{6}$$

The equivalent nominal bid–ask spread is

$$s_t^L \equiv \mathcal{S}_t^L p_t = \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L} p_t. \tag{7}$$

Proposition 1. (Continuous pricing bid–ask spread) *With zero tick size and lot size L, the equilibrium percentage spread is* $\mathcal{S}_t^L = \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L}$ *and the equilibrium bid–ask spread is* $s_t^L = \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L} p_t$*:*

(i) The market maker provides exactly one lot of liquidity at $\mathrm{p}_t \pm \frac{s_t^L}{2}$ *and she refills one lot around* p_t *after each trade.*

(ii) Uninformed traders slice their demand into a series of child orders of one lot each.

(iii) Informed traders adversely select one lot of liquidity per jump, and p_t *updates afterwards.*

Our model degenerates to BCS when $h = 1$, $L = 1$, and $\Delta = 0$ and when v_t is public information. The percentage spread converges to 0 when $L \to 0$ or $h \to \infty$. Therefore, our paper offers two possible solutions to the sniping problem in BCS. The policy solution is to reduce the lot size and the market solution is aggressive stock splits. When prices are continuous, the policy and market solutions are economically equivalent because both reduce lot sizes in dollars. In Sections 4 and 5, we show that policy and market solutions are no longer equivalent when pricing becomes discrete, because lot-size reduction makes quantity more continuous without affecting price discreteness, but stock splits make quantity more continuous while making the price more discrete.

Multiply both the denominator and the numerator of Equation (7) by p_t, and we have

$$s_t^L = \frac{2\sigma\lambda_J L p_t^2}{\lambda_I p_t h + \lambda_J p_t L}. \tag{8}$$

Notice that the denominator equals the dollar volume per unit of time. Denote

$$DVol_t \equiv \lambda_I p_t h + \lambda_J p_t L = \lambda_I v_t + \lambda_J p_t L. \tag{9}$$

We then discover the following Square Rule for the bid–ask spread:

Corollary 1 (Square Rule). *Under continuous pricing, the nominal bid–ask spread $s_t^L = \frac{2\sigma\lambda_J L}{DVol_t} p_t^2$ is proportional to the square of the nominal price, controlling for dollar trading volume and stock volatility.*

Corollary 1 shows that an increase in the nominal price leads to a quadratic increase in the bid–ask spread. A p_t-time increase comes from the linear increase in the price, leaving the percentage spread unchanged. Another p_t-time increase comes from the increase in adverse-selection risk, as the cost to sustain one-round-lot liquidity also increases linearly in the price. Combining the two effects, we have $s_t^L \propto p_t^2$.

3.2 The Firm's Choice

The firm aims to minimize its expected transaction cost by choosing h.[8] The firm's dollar volume per unit of time is $\lambda_I p_t h + \lambda_J p_t L$, and the firm's percentage spread is $S_t^L = \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L}$. Therefore, the firm's objective function is

$$\min_h \mathbb{E}\left[\left(\lambda_I p_t h + \lambda_J p_t L\right) \cdot \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L}\right] = \mathbb{E}\left[\sigma\lambda_J L p_t\right] = \sigma\lambda_J L \cdot \mathbb{E}[p_t] = \sigma\lambda_J L p \equiv \sigma\lambda_J L \cdot \frac{v}{h}. \tag{10}$$

The firm's objective function under continuous pricing is highly intuitive. $\sigma\lambda_J L \cdot \frac{v}{h}$ is

[8] Notice that minimizing the expected transaction cost is economically equivalent to minimizing the expected percentage spread. In fact, when the price is continuous, the minimization problems are also mathematically equivalent. We make a technical assumption that the firm minimizes the expected total transaction cost because, when the tick size does not equal zero, the percentage spread includes a term $\frac{\Delta}{p_t}$. $\mathbb{E}\left(\frac{\Delta}{p_t}\right)$ does not have an analytical form because the denominator is a random variable.

the market maker's expected adverse-selection cost per unit of time. Thus, the firm's objective function is equivalent to minimizing its market maker's adverse-selection cost. As the market maker must maintain one round lot of liquidity, a decrease in either L or p reduces dollar lot size pL and thereby the market maker's adverse selection costs. The market maker can still accommodate demand for liquidity through refilling the liquidity more frequently at the smaller dollar lot size.

Under continuous pricing, firms should choose $h \to \infty$ and $p \to 0$. The result is intuitive. When lot size is the only friction, firms should split their stocks aggressively to minimize the friction from discrete quantities. The constraint that prevents the firm from choosing very low prices comes from the other friction: discrete pricing. We consider the tradeoff between discrete pricing and discrete quantities in the next section.

4. DISCRETE PRICING AND DISCRETE LOTS

We present our main theoretical predictions in this section, where the regulator chooses a discrete tick size Δ such that trades and quotes can occur only at the pricing grid $\{\Delta, 2\Delta, 3\Delta, \cdots\}$. As a firm cannot reduce its bid–ask spread below one tick, splits increase frictions generated by discrete prices and may increase expected transaction costs. The tick constraint, therefore, favors high prices. We solve the model through backward induction. In Subsection 4.1, we quantify the frictions generated by the discrete tick size and solve traders' optimal decisions given h, L, and Δ. In Subsection 4.2, the firm solves the optimal nominal shares outstanding h, balancing the frictions generated by lot and tick sizes. Subsection 4.3 presents the formula for the optimal split ratio.

4.1. Traders' Decisions and Friction from the Tick Size

Under discrete pricing, the market maker can no longer quote competitive prices at $p_t \pm \frac{s_t^L}{2}$. Lemma 1 shows that she quotes a bid price at the tick immediately below $p_t - \frac{s_t^L}{2}$ and an ask price at the tick immediately above $p_t + \frac{s_t^L}{2}$.

Lemma 1 (Discrete Pricing Bid–ask Spread). *With tick size* Δ, *the competitive market maker quotes an ask price,* $A_t = p_t + \frac{s_t^L}{2} + \left[\Delta - mod\left(p_t + \frac{s_t^L}{2}, \Delta\right)\right]$, *and a bid price,* $B_t = p_t - \frac{s_t^L}{2} - \left[\Delta - mod\left(p_t - \frac{s_t^L}{2}, \Delta\right)\right]$, *where* $mod(x, y) \in [0, y)$ *is the remainder of dividing x by y.*

Lemma 1 shows that the competitive market maker rounds up the continuous ask price to the tick immediately above and rounds down the bid price to the tick immediately below, because more aggressive quotes lose money while less aggressive quotes are not competitive. Therefore, the quotes are $\left[\Delta - mod\left(p_t + \frac{s_t^L}{2}, \Delta\right)\right]$ higher on the ask side and $\left[\Delta - mod\left(p_t - \frac{s_t^L}{2}, \Delta\right)\right]$ lower on the bid side (the green arrows in Figure 1).[9]

Next, we calculate the widening effect for any break-even spread s_t^L. First, we can decompose the break-even spread into two components, $s_t^L = a\Delta + b$, where $a = 0,1,2,3, \ldots$ and $b = mod(s_t^L, \Delta)$. The first component is rendered in the multiple ticks and the second component is the residual that is narrower than one tick. Proposition 2 shows that the widening effect is either $\Delta - b$ or $2\Delta - b$ depending on the relative position of p_t within the tick grids, $mod(p_t, \Delta)$. More importantly, Proposition 2 shows that the expectation of the widening effect is one tick.

[9] The tick size creates rents for the competitive market maker. With the artificially wide bid–ask spread, the depth at the NBBO might be greater than 1 round lot (Yao and Ye 2018). The exact depth at time t depends on the position of p_t within the tick grid and the break-even spread s_t^L. We do not characterize the depth in this model because the depth does not affect the bid–ask spread. In equilibrium, the depth becomes one round lot when $p_t \pm \frac{s_t^L}{2}$ is close to the tick grid, and it can be greater than one round lot when $p_t \pm \frac{s_t^L}{2}$ is far away from the tick grid (see Li, Wang, and Ye 2021 for details). The uninformed traders' optimal strategy is to choose $q = 1$ because they need to decide their λ_q before trading starts. If they choose a greater order size, they'll walk the book when the depth is only one round lot, which incurs greater transaction costs.

Proposition 2 (Average Widening Effect) *Define the bid–ask spread under discrete prices as* $s_t^{tot} = B_t - A_t$ *and define the widening effect at any time* t *as* $s_t^{\Delta} = s_t^{tot} - s_t^L$, where s_t^L is the *competitive bid–ask spread under continuous pricing. We have*

i) $$s_t^{\Delta} = \begin{cases} \Delta - b, \ if \begin{cases} mod(p_t, \Delta) \in \left[\frac{b}{2}, \Delta - \frac{b}{2}\right] \text{ and } a \ is \ even \\ mod(p_t, \Delta) \in \left[\frac{\Delta}{2} + \frac{b}{2}, \Delta\right) \cup \left[0, \frac{\Delta}{2} - \frac{b}{2}\right] \text{ and } a \ is \ odd \end{cases} \\ 2\Delta - b, \ if \begin{cases} mod(p_t, \Delta) \in \left[0, \frac{b}{2}\right) \cup \left(\Delta - \frac{b}{2}, \Delta\right) \text{ and } a \ is \ even \\ mod(p_t, \Delta) \in \left(\frac{\Delta}{2} - \frac{b}{2}, \frac{\Delta}{2} + \frac{b}{2}\right) \text{ and } a \ is \ odd \end{cases} \end{cases}$$

ii) *If* $p_t \gg \Delta$, $mod(p_t, \Delta) \xrightarrow{d} U[0, \Delta)$ *and* $\mathbb{E}\left(s_t^{\Delta}\right) = \Delta$.

The intuition for the one-tick average widening effect is as follows. For a Poisson jump process, p_t follows a lognormal distribution. As the distribution is smooth, the residual of p_t should not cluster at any specific position within the tick. As any residual value within the tick is equally likely, $mod(p_t, \Delta)$ converges to a uniform distribution as t grows. This argument is generally true for any distribution with a smooth, non-clustering probability density function.[10]

When the residual of p_t is uniformly distributed, the average widening effect is one tick in magnitude for any break-even bid–ask spread. Figure 1 presents the intuition for this result using a small break-even spread of 0.2 ticks and a large break-even spread of 0.8 ticks. In this example, $a = 0$, but the intuition holds for any a. Yellow dotted lines represent the tick grids, red (blue) upper solid lines represent the continuous pricing ask (bid) prices, and green arrows represent the widening effects. In the two panels on the left,

[10] The only exception is when $\frac{p_t}{\Delta} \to 0$, i.e. when the nominal price of the stock is too low. In this case, $mod(p_t, \Delta)$ may cluster around 0. In the equilibrium of our model, the firm should not choose such a low price because it suffers from dramatic tick-size constraints. In reality, both the NYSE and NASDAQ delist a stock if its price falls under \$1.00 (i.e. when $p_t < 100\Delta$). Therefore, $p_t \gg \Delta$ generally holds. In the same spirit, Anshuman and Kalay (1998) assume that p_t follows a normal distribution whose variance increases in t. When t is large enough, the standard deviation of p_t becomes much larger than the tick size, and p_t is asymptotically uniformly distributed within the tick grids.

p_t is "lucky" because the break-even bid and ask prices are at the same tick grid. Therefore, the widening effect is less than one tick ($s_t^\Delta = \Delta - b$). In the two panels on the right, p_t is "unlucky" because the break-even bid and ask prices are at different tick grids. The widening effect is then more than one tick ($s_t^\Delta = 2\Delta - b$).

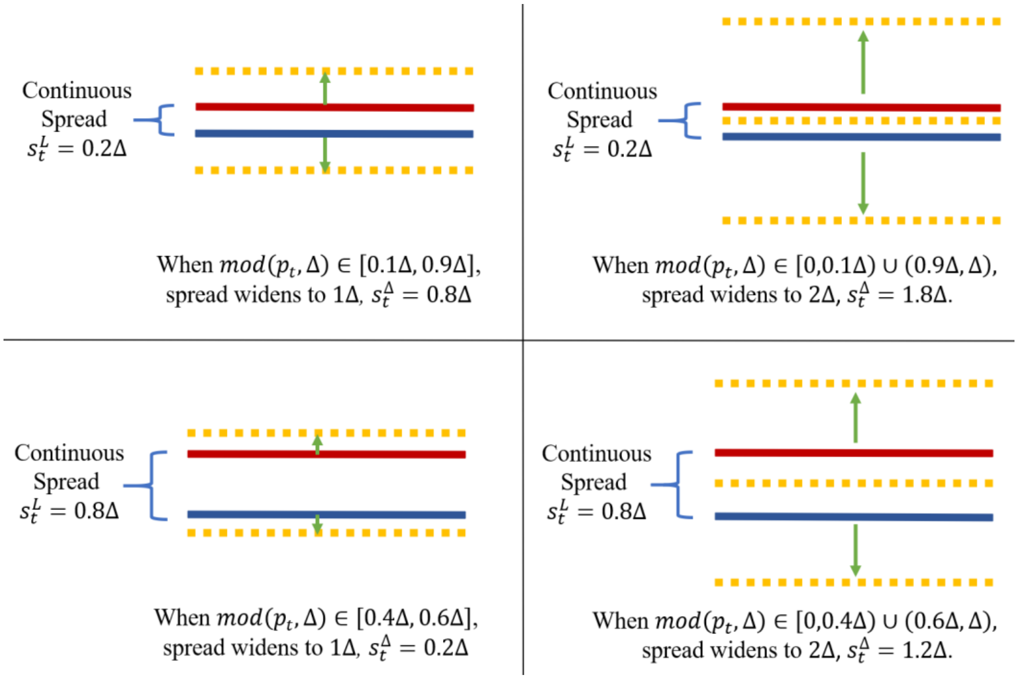

FIGURE 1. —Average widening effect of one tick: In this figure we illustrate the bid–ask widening effect. The red solid upper bars are the ask prices under continuous pricing and the blue solid lower bars are the bid prices under continuous pricing. The yellow dots represent the tick grids. The green arrows are the widening effects, where the ask prices move up to the next available tick grid and the bid prices move down to the next available tick grid. The two panels on the left illustrate the "lucky" cases where the bid–ask spread widens by less than one tick, and the two panels on the right illustrate the "unlucky" cases where the bid–ask spread widens by more than one tick.

Proposition 2 shows that the expected widening effect for any break-even spread is one tick. Figure 1 provides the intuition underlying this result using $a = 0$ as an example. When b is small, the best bid and ask are more likely to be within the same tick such that the widening effect is $\Delta - b$ but not $2\Delta - b$. However, a small b increases $\Delta - b$ and $2\Delta - b$. The probability effect and the widening effect cancel out exactly when p_t is uniformly distributed within the tick. For example, for a small $b = 0.2\Delta$, $p_t \pm \frac{s_t^L}{2}$ are within the same tick 80% of the time, but the widening effect is either 0.8 ticks or 1.8 ticks. For a large $b =$

0.8Δ, $p_t \pm \frac{s_t^L}{2}$ are within the same tick only 20% of the time, but the widening effect is either 0.2 ticks or 1.2 ticks. In general, p_t is lucky with probability $\frac{\Delta-b}{\Delta}$ and unlucky with probability $\frac{b}{\Delta}$ for any a. Therefore, the expectation for the widening effect is

$$\mathbb{E}\left(s_t^\Delta\right) = \frac{\Delta-b}{\Delta} \cdot (\Delta - b) + \frac{b}{\Delta} \cdot (2\Delta - b) = \Delta. \tag{11}$$

Equation (11) implies that we can decompose the average bid–ask spread under discrete pricing, $\mathbb{E}(s_t^{tot})$, into a lot-driven component $\mathbb{E}(s_t^L)$ and a tick-driven component $\mathbb{E}\left(s_t^\Delta\right) = \Delta$. This decomposition allows us to quantify the friction from discrete price and discrete quantity to the percentage spread. A uniform widening effect of one tick implies that the tick friction is proportional to p^{-1}. The quadratic relationship between the lot-driven spread and the price implies that the lot friction is proportional to p. Therefore, an increase in p reduces tick-size friction and increases lot-size friction. In the next subsection, the firm chooses the optimal p, which minimizes the sum of these two frictions.

4.2. The Firm's Decision and the Optimal Nominal Price

The firm chooses h (and equivalently $p \equiv \frac{v}{h}$) to minimize the expected execution cost, given lot size L and tick size Δ. The firm's objective function is

$$\begin{aligned}
\min_p \mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) &= \min_p \mathbb{E}\left(\frac{s_t^L + s_t^\Delta}{2p_t} \cdot DVol_t\right) \\
&= \min_p \mathbb{E}\left[\left(\frac{\frac{2\sigma\lambda_J L p_t}{\lambda_I h + \lambda_J L} + s_t^\Delta}{2p_t}\right) \cdot \left(\lambda_I p_t h + \lambda_J p_t L\right)\right] \\
&= \min_p \mathbb{E}\left[\sigma\lambda_J p_t L + \frac{s_t^\Delta}{2}\left(\lambda_I h + \lambda_J L\right)\right] \\
&= \min_p \left[\sigma\lambda_J L \cdot \mathbb{E}[p_t] + \frac{\mathbb{E}(s_t^\Delta)}{2}\left(\lambda_I h + \lambda_J L\right)\right] \\
&= \min_p \left[\sigma\lambda_J L p + \frac{\Delta}{2}\lambda_I \frac{v}{p} + \frac{\Delta}{2}\lambda_J L\right]
\end{aligned} \tag{12}$$

The first term in the last line ($\sigma\lambda_J Lp$) measures the expected execution cost that is driven by the lot size. The second term ($\frac{\Delta}{2}\lambda_I\frac{v}{p}$) measures the expected execution cost that is driven by the tick size. The third term is a constant. Applying the inequality of arithmetic and geometric means, we have

$$\sigma\lambda_J Lp + \frac{\Delta}{2}\lambda_I\frac{v}{p} \geq 2\sqrt{\sigma\lambda_J Lp \cdot \frac{\Delta}{2}\lambda_I\frac{v}{p}}. \tag{13}$$

The equality holds only when $\lambda_J Lp = \frac{\Delta}{2}\lambda_I\frac{v}{p}$, or when the impact of the lot size is equal to the impact of the tick size. The corresponding optimal nominal price is

$$p^* = \sqrt{\frac{\lambda_I \Delta v}{2\sigma\lambda_J L}}. \tag{14}$$

Proposition 3 (Golden Rule of Two Cents). *When the tick size is* Δ *and the lot size is* L*, the optimal nominal price is* $p^* = \sqrt{\frac{\lambda_I \Delta v}{2\sigma\lambda_J L}}$ *and* $\mathbb{E}(s^{tot}) = \Delta \cdot \left(1 + \frac{\lambda_I h}{\lambda_I h + \lambda_J L}\right)$. *When* $h \gg L$, $\mathbb{E}(s^{tot}) \approx 2\Delta$.

Surprisingly, Proposition 3 shows that all firms reach their optimal price when their nominal bid–ask spread is two ticks wide.[11] The intuition for the uniform optimal bid–ask spread is as follows. As the tick size creates friction of one tick for all stocks, firms should manage their prices such that the friction from the lot is also equal to one tick. Stocks whose bid–ask spreads are narrower than two ticks are constrained to a greater extent by the tick size, and those firms should increase their nominal prices. Stocks whose average bid–ask spreads are wider than two ticks are constrained to a greater extent by lot size, and those firms should decrease their prices.

[11] This result holds under the trivial assumption that each firm's shares outstanding h is much larger than its lot size L. The minimal shares outstanding in our sample is 56,600 round lots.

The homogeneous two-tick spreads then lead to heterogeneous nominal prices. Any economic force that reduces the percentage spread would increase the optimal price to maintain the two-tick optimal spread. An increase in volatility, caused either by an increase in jump size σ or by an increase in jump frequency λ_J, increases the market maker's adverse selection cost and percentage spread. Therefore, we predict that volatile stocks will choose lower prices. An increase in dollar volume, caused either by an increase in market cap v or by an increase in turnover rate λ_I, provides more revenue to market makers and reduces the percentage spread. Therefore, we predict that active stocks will choose higher prices.

We can also understand the comparative statics of optimal prices using lot- or tick-size constraints. Take volatility as an example. Holding all else equal, an increase in volatility increases adverse-selection risk and intensifies lot constraints. Therefore, the firm should reduce the price until the lot constraint is again equal to the tick constraint. We can also consider that an increase in volatility makes the movement of the price and bid–ask spread constrained to a lesser extent by the tick size, giving the firm a greater incentive to choose a lower price.

4.3. Optimal Split Ratio: The Modified Square Rule

Our parsimonious model also provides a simple formula enabling firms to adjust to their optimal prices. Firms do not need to calibrate σ, λ_J, and λ_I to estimate the optimal price because Corollary 2 shows that the average bid–ask spread provides a sufficient statistic for firms to choose their optimal split ratios.

Recall that Proposition 2 shows that s_t^{tot} can be decomposed into two components: s_t^L and s_t^Δ. The expectation of the tick-driven component $\mathbb{E}(s_t^\Delta)$ is always Δ. Therefore

$$\mathbb{E}(s_t^{tot}) = \mathbb{E}\left(s_t^L + s_t^\Delta\right) = \mathbb{E}(s_t^L) + \Delta. \tag{15}$$

Equation (7) shows that $s_t^L = p_t S_t^L = \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L} p_t$, and

$$\mathbb{E}(s_t^L) = \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L} p = \frac{2\sigma\lambda_J L}{\lambda_I p h + \lambda_J p L} p^2 = \frac{2\sigma\lambda_J L}{\lambda_I p h (1 + \frac{\lambda_J L}{\lambda_I h})} p^2 = \frac{2\sigma\lambda_J L}{\lambda_I v (1 + \frac{\lambda_J L}{\lambda_I h})} p^2 \approx \frac{2\sigma\lambda_J L}{\lambda_I v} p^2. \quad (16)$$

Again, when shares outstanding $h \gg L$, the expectation of a lot-driven spread increases in the square of p^2. Therefore, an H-for-1 split reduces the expected lot-driven component by H^2-fold, or $\frac{\mathbb{E}(s_t^{tot})-\Delta}{H^2}$. As the expectation of the tick-driven component is always Δ, the expected bid–ask spread post-split is

$$\mathbb{E}(s_t^{tot,post}) = \frac{\mathbb{E}(s_t^{tot})-\Delta}{H^2} + \Delta. \quad (17)$$

We call formula (17) the Modified Square Rule.

Corollary 2 (The Modified Square Rule). *When $h \gg L$, an H-for-1 split changes the spread from $\mathbb{E}(s_t^{tot})$ to $\frac{\mathbb{E}(s_t^{tot})-\Delta}{H^2} + \Delta$.*

The Modified Square Rule predicts post-split changes in bid–ask spreads. We show in Section 8.2 that any predicted change matches up almost one for one with the realized change. The Modified Square Rule explains why firms with similar fundamentals can have dramatically different liquidity. For example, Amazon's stock was priced at $3,305 per share and its bid–ask spread was 153 cents (4.62 bps), while Microsoft's stock was priced at $255 per share and its bid–ask spread was 1.95 cents (0.77 bps). We find that the difference in nominal prices almost fully explains the sixfold difference in transaction costs. If Amazon were to split 13-for-1, the Modified Square Rule predicts a new nominal spread of $\frac{153-1}{13^2} + 1 = 1.90$ cents (0.75 bps), which is similar to Microsoft's spread.

The Modified Square Rule provides a convenient tool that firms can use to find their optimal prices. The split ratio H that achieves the optimal two-tick spread satisfies $\frac{\mathbb{E}(s_t^{tot})-\Delta}{H^2} + \Delta = 2\Delta$. The solution is $H^* = \sqrt{\frac{\mathbb{E}(s_t^{tot})-\Delta}{\Delta}}$.

Corollary 3 (The Optimal Split Ratio). *When $h \gg L$, the split ratio to achieve the liquidity-optimal price is* $H^* = \sqrt{\frac{\mathbb{E}(s_t^{tot})-\Delta}{\Delta}}$.

For example, the liquidity-optimal split ratio for Amazon is $\sqrt{\frac{153-1}{1}} = 12.3$.[12] The optimal split ratio would further reduce Amazon's percentage spread to 0.74 bps. We use the optimal split ratio in our back-of-the-envelope calculation for the liquidity benefit to adjust to the optimal price.

5. POLICY IMPLICATIONS FOR TICK AND LOT SIZES

In this section, we allow the regulator to change tick and lot sizes. In Subsection 5.1, we allow the regulator to increase or decrease the uniform tick and lot sizes. In Subsection 5.2, we allow the regulator to switch to proportional tick and lot sizes, in which case the tick and lot sizes are functions of price.

5.1. Tick and Lot Size Changes under a Uniform System

We find that an increase in the tick or lot size reduces liquidity. The firm can adjust its optimal price to mitigate the negative shocks, but only partially. Corollary 4 shows that the firm's best response and the change in liquidity both follow the square root of the change in the tick and lot sizes.

Corollary 4. (Square Root Rule) *A firm's optimal nominal price* $p^* = \sqrt{\frac{\lambda_I v \Delta}{2\sigma\lambda_J L}}$ *responds to tick- and lot-size changes by* $\sqrt{\Delta/L}$. *When* $h \gg L$ *, the average nominal spread*

[12] After we completed the first draft of our paper, Amazon indeed announced a 20-for-1 stock split, on March 9, 2022. We predicted that the split would dramatically reduce the percentage bid–ask spread for Amazon. Because the split ratio was greater than 12.3-for-1, we predicted that the spread would be lower than two cents after the split implementation date (June 3, 2022). Consistent with our prediction, the bid-ask spread of Amazon is now around 1.7 ticks (1.5 bps), a 68% reduction from the 4.62 bps of pre-split spread.

under optimal pricing equals 2Δ *regardless of* L *and* Δ*. The smallest achievable execution cost* $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) \approx \sqrt{2\sigma v \lambda_I \lambda_J L\Delta}$ *is proportional to* $\sqrt{L\Delta}$*.*

The comparative statics of p^* show that a firm's optimal response to a change in the tick or lot size is found in its square roots. For example, if regulators increase the tick size from one cent to five cents, firms should reverse-split their stocks by $\sqrt{5}$.[13] This optimal reverse-split ratio increases the relative tick size and dollar lot size by the same proportion, such that the marginal contribution of the tick size is still equal to the marginal contribution of the lot size.

Angel (1997) considers only discrete pricing. In his framework, a 1-for-5 reverse split would neutralize a fivefold increase in the tick size. When we add discrete quantities, a 1-for-5 reverse split neither neutralizes the increase in the tick size nor is the best response. In fact, Corollary 4 shows that a 1-for-5 reverse split leads to the same transaction costs as doing nothing at all. Although a 1-for-5 reverse split restores the relative tick size, such aggressive reverse splits cause a fivefold increase in the dollar lot size. As Corollary 4 shows, the total friction from the tick and lot sizes is proportional to $\sqrt{L\Delta}$, so a fivefold increase in the dollar lot size leads to the same increase in the transaction cost as a fivefold increase in the tick size.

For example, consider a firm currently at its optimal spread of two cents, one cent from the tick size and one cent from the lot size. An increase in the tick size from one cent to five cents raises the tick-driven spread to five cents, leading to a nominal spread of six cents, which is three times the previous level. After a 1-for-5 reverse split, the tick-driven spread remains at five cents. A fivefold increase in the lot size raises the lot-driven spread to 5^2. The nominal spread now becomes $5 + 5^2$. After adjusting for the fivefold increase in the nominal price, the transaction cost still increases by a factor of three ($= \frac{25+5}{2\times 5}$). In conclusion, a reverse split at the same rate as the tick size increases is equivalent to doing

[13] Here we consider a permanent fivefold tick-size increase. The U.S. Tick Size Pilot program increased the tick size temporarily for two years. We show in Section 8.4 that the fixed costs of splits may outweigh the benefits for first reverse-splitting and then splitting back to counteract a two-year temporary shock.

nothing at all.

Corollary 4 shows that the new optimal price can neutralize the change in the tick or lot size, but only partially. In the new equilibrium, transaction costs also change at the rate of the square root. To see this, recall that the optimal 1-for-$\sqrt{5}$ reverse split increases the lot-driven spread to 5 ($= \sqrt{5} \times \sqrt{5}$) cents while the tick-driven spread remains 5 cents. The 1-for-$\sqrt{5}$ reverse split restores the two-tick optimal spread, except that the two ticks now equal ten cents. The optimal bid–ask spread increases fivefold and the nominal price increases by a factor of $\sqrt{5}$, leading to a $\sqrt{5}$-fold increase in transaction costs. In summary, the Two-Tick Rule always holds, but the firm's optimal price changes in accordance with the new tick size.

The same intuition applies to a reduction in the lot size. In 2019, the SIP Operating Committee solicited comments for a policy initiative designed to reduce the friction associated with odd-lot trades, or orders involving fewer than 100 shares. Stock exchanges and institutional traders proposed a more aggressive plan: reduce the round-lot threshold to fewer than 100 shares.[14] Corollary 4 indicates that a reduction in the lot size improves liquidity and that firms should reverse-split their stocks to take full advantage of this benefit. For example, if the SIP committee were to reduce the round lot from 100 shares to 1 share, firms should reverse-split at a ratio of 1-to-$\sqrt{100}$ to maximize the benefit of the lot-size reduction. Such a reduction in the spread would also explain why broker-dealers, who often provide execution within the bid–ask spread against retail traders (Boehmer et al., 2020), oppose any reduction in the official lot size.[15] A reduction in the lot size narrows the reference bid–ask spread in stock exchanges and thereby forces these brokers to offer better prices to retail traders.

One of the most vexing puzzles in the literature on nominal prices and the tick size involves understanding why firms did not split 1-to-6.25 after the 2001 decimalization

[14] NASDAQ's comment letter pertaining to this plan suggests that "high value quotations with significant price discovery information would be protected, even if they were less than 100 shares." Citadel and Blackrock also support lot-size reduction in their comment letters. Their comment letters can be found at https://www.ctaplan.com/oddlots.

[15]. The opposition from retail broker-dealers can also be found at https://www.ctaplan.com/oddlots.

standard reduced the tick size by a factor of 6.25 (Weld et al., 2009). Our Square Root Rule first reduces the gap from 6.25 to $\sqrt{6.25} = 2.5$. The market crash of 2001 may then further fill the gap around decimalization. Most importantly, our model predicts that the optimal price will be the same if the tick size and lot size are reduced by the same amount. The proliferation of electronic trading allows traders to slice their orders into smaller pieces and effectively reduce lot sizes. Electronic trading provides an incentive to choose high prices, which counteracts the incentive to reduce prices led by decimalization. Also, as decimalization is a one-time shock and electronic trading evolves over the years, we see an increase in nominal prices and a lack of stock splits over the past two decades (Mackintosh 2021).

5.2 Proportional vs. Uniform Tick and Lot Sizes

One plan for changing lot size is to make it a function of price, such that high-priced stocks have smaller lot sizes.[16] This plan generates a proportional lot size, leading to a more uniform dollar lot size for all stocks. Also, in many European countries, Hong Kong, and Japan, the tick size increases with stock prices. Corollary 5 shows that if the tick size is proportional, firms should split their stocks to minimize friction driven by the lot size. On the other hand, if the lot size is proportional, firms should reverse-split to minimize the tick-size friction. If both the lot and tick size are proportional, the choice of a nominal price becomes irrelevant.

Corollary 5. (Proportional Tick and Lot Systems) *(1) With fixed Δ and proportional lot size $\mathbb{L}(p) = k^L/p$, where k^L is a constant, the firm's optimal choice is $p^* \to \infty$ and*

[16] For example, Blackrock (2019) "believes that a more elegant solution for the inclusion of odd lots would be to move from a 'one-size-fits-all' approach to a multi-tiered framework where round lot sizes are determined by the price of a security." NASDAQ (2019) suggests "establishing a dollar threshold for the value of quotes to be protected, defined as price multiplied by the number of shares." Consequentially, the SEC proposes a 3-tier lot-size regime for U.S. equities: 1 share for stocks priced at $10,000.01 or more per share, 10 shares for stocks priced $1,000.01 to $10,000.00 per share, 40 shares for stocks priced $250.01 to $1,000.00 per share, and 100 shares for other stocks.

$\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) = \sigma\lambda_J k^L$.*(2) With fixed L and proportional tick size* $\Delta(p) = k^\Delta p$, *where* k^Δ *is a constant, the firm's optimal choice is* $p^* \to 0$ *and* $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) = \frac{k^\Delta \lambda_I v}{2}$. *(3) With proportional tick* $\Delta(p) = k^\Delta p$ *and lot* $\mathbb{L}(p) = k^L/p$, $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) \equiv \sigma\lambda_J k^L + \frac{k^\Delta \lambda_I v}{2}$ *for any* p. *Adopting this proportional system with any reference price* p_Ω *such that* $k^\Delta = \Delta/p_\Omega$ *and* $k^L = Lp_\Omega$ *reduces liquidity for any stock with* $p \neq p_\Omega$.

In Table 1, we summarize the results derived from Corollary 5. The intuitions are as follows. Suppose the regulator chooses lot size $\mathbb{L}(p)$ and tick size $\Delta(p)$. A firm that chooses a price p then has a dollar lot size $p\mathbb{L}(p)$ and a relative tick size $\frac{\Delta(p)}{p}$. Given $p\mathbb{L}(p)$ and $\frac{\Delta(p)}{p}$, the game in stage 3 is like the trading game under uniform tick and lot sizes. The lot-driven spread still follows the Square Rule, except that it increases in the square of $p\mathbb{L}(p)$; the tick-driven spread is still one tick, except that one tick is now $\Delta(p)$.

TABLE 1
Optimal Price with Fixed and Proportional Tick/Lot Sizes

Tick Size \ Lot Size	Fixed L	Proportional $\mathbb{L}(p) = k^L/p$
Fixed Δ	$p^* = \sqrt{\frac{\lambda_I v \Delta}{2\sigma\lambda_J L}}$	$p^* \to \infty$
Proportional $\Delta(p) = k^\Delta p$	$p^* \to 0$	Expected transaction cost does not depend on p

In this table we summarize the firm's optimal choices regarding price p^* under various tick- and lot-size systems. Continuous pricing is a special case for the first row, where $\Delta = 0$, and continuous quantities is a special case for the first column, where $L = 0$. Proportional tick and lot sizes are summarized in the second row and column, respectively. Reads: With fixed tick and lot sizes, firms choose the optimal nominal price p^*. Firms split (reverse-split) to the extreme if the tick (lot) size is proportional to the nominal price. If both tick and lot sizes are proportional, the nominal price no longer affects the firm's liquidity.

If one variable is uniform and the other is proportional, the firm's optimal choice is to minimize the friction caused by the uniform variable. If the tick size is proportional but the lot size is uniform, stock splits do not change the proportional tick size $\frac{\Delta(p)}{p}$ but reduce the dollar lot size. Therefore, the firm should split aggressively to reduce the lot-driven transaction cost. If the tick size is uniform but the lot size is proportional, reverse splits do not increase the dollar lot size but reduce the relative tick size. Therefore, the firm should choose a high price to minimize the tick-driven transaction cost.

When tick and lot sizes are both proportional, a change in price changes neither relative tick size nor dollar lot size. Whether the move from the uniform tick and lot sizes to proportional tick and lot sizes improves liquidity depends on the relative tick size (k^{Δ}) and dollar lot size (k^{L}) mandated by the regulator. A natural way to choose k^{Δ} and k^{L} is to use a representative stock. For example, a regulator can choose k^{Δ} and k^{L} such that the relative tick size and dollar lot sizes for a $30 benchmark stock do not change. Corollary 5 shows that such proportional systems reduce liquidity for *all* stocks except the benchmark. The greater the distance between the stock price and the benchmark price, the greater the liquidity reduction.

For example, a proportional system that retains the relative tick and dollar lot sizes for a $30 stock would impose a tenfold wider tick size and a 0.1-fold larger lot size on a $300 stock. Suppose the $300 stock previously had a two-cent optimal bid–ask spread, one cent from the tick constraint and one cent from the lot constraint. The move to the proportional system would increase its tick-driven spread to ten cents and reduce its lot-driven spread to 0.1 cents, leading to an increase in the total spread from two cents to 10.1 cents.[17] Symmetrically, the proportional system would impose a 0.1-fold larger tick size and tenfold larger lot size for a $3 stock. If the $3 stock currently trades with a two-cent bid–ask spread, its tick-driven spread would drop to 0.1 cents, but its lot-driven spread would increase to 10 cents. The total spread is again 10.1 cents. Under uniform tick and lot sizes, a firm choosing a $300 ($3) price is more (less) liquid than a firm choosing a $30 price, but adopting a proportional tick and lot system reduces liquidity for both the $300 and the

---

[17] The change is linear in the change in lot size because the regulator does not change the price of the stock. Under the Square Rule, the firm increases both its dollar lot size and the price, leading to a quadratic relationship.

\$3 stocks by the same magnitude. Corollary 5 implies that regulators should not use any existing stock as the benchmark if they want to switch from a uniform system to a proportional system. [18]

The uniform system may seem less flexible because it mandates the same tick and lot sizes for stocks listed at varying prices. Yet the uniform system gives firms the flexibility to choose the optimal balance between lot and tick sizes by adjusting nominal prices. More liquid stocks endogenously choose higher prices (i.e., larger dollar lot sizes and smaller relative tick sizes) because the main friction comes from discrete pricing. Less liquid stocks endogenously choose lower prices (i.e., smaller dollar lot sizes and larger relative tick sizes) because the main friction comes from trading large lots. The proportional system is actually less flexible because it mandates the same level of price and quantity discreteness for firms with varying fundamentals. The biggest victims would be stocks whose optimal nominal prices (and implicitly their stock characteristics) differ to the greatest extent from the benchmark stock.

As our paper focuses on frictions caused by tick and lot sizes, the first best in our model is continuous tick and lot sizes. This result is consistent with Kyle and Lee (2017) and Budish et al. (2022), who propose a fully continuous exchange. Uniform tick and lot sizes offer one degree of freedom enabling firms to balance discrete pricing and quantities by choosing the nominal price. Proportional tick and lot sizes offer zero degrees of freedom because they mandate the same level of discreteness in price and quantity for stocks with heterogeneous characteristics.

6. A THREE-VARIABLE EMPIRICAL MODEL OF LIQUIDITY

A fundamental question for market microstructure involves identifying the

[18] The proportional plan in this example harms liquidity because it increases lot size for low-priced stocks. The SEC's plan reported in footnote 15 reduces lot size for high-priced stocks and does not increase the lot size for any stocks. Therefore, our model predicts the liquidity would improve under the SEC's plan. .

determinants of bid–ask spreads (Stoll 2000). In this section, we show that our three-variable model of liquidity explains 81% of the cross-sectional variation in bid–ask spreads. Surprisingly, our model has a much higher R^2 than existing benchmarks (Madhavan 2000; Stoll 2000) even though we use only a subset of their variables.

Corollary 1 decomposes the bid–ask spread into a tick-driven component and a lot-driven component:

$$s_t^{tot} - s_t^{\Delta} = s_t^L = \frac{2\sigma\lambda_J L}{DVol_t} p_t^2. \tag{18}$$

Taking the natural log on both sides, we obtain

$$log(s_t^{tot} - s_t^{\Delta}) = 2log(p_t) - log(DVol_t) + log(\sigma\lambda_J) + const. \tag{19}$$

Following the literature, our variable of interest is the time-weighted average spread. As our horizon is one year, it is safe to assume that p_t has evolved over a long enough period such that the average widening effect $E(s_t^{\Delta})$ becomes Δ, following Proposition 2. We use the stock volatility, i.e., standard deviation of daily stock returns, as the empirical proxy for $\sigma\lambda_J$. Therefore, we can write (19) in the form of an OLS test:

$$log\left(\overline{Spread} - \Delta\right)_i = \alpha + \delta \cdot log(\overline{Price})_i + log\left(\overline{Volume}\right)_i + log\left(\overline{Volatility}\right)_i + \varepsilon_i. \tag{20}$$

Our Modified Square Rule predicts that $\delta = 2$.[19] The null hypothesis is $\delta = 1$: when lot size does not impose a binding constraint on the bid–ask spread, $\overline{Spread} - \Delta$ should increase one-to-one in price.[20]

Our sample includes all U.S.-listed common stocks (SHRCD 10 or 11) with a standard lot size of 100 shares and a standard tick size of 1 cent.[21] Our main sample period is the year 2020, and we conduct robustness checks using previous years. We use daily Trade

[19] Equation (20) presents the Modified Square Rule in terms of price. The result follows the same economic mechanism as in Corollary 2, where we present a similar result using the split ratio.

[20] Consider two stocks that are identical except for their nominal prices. If lot size does not impose any constraint on order size, investors should submit orders of the same dollar size for both stocks. The market maker then displays the same dollar amount of liquidity, and the break-even percentage spread is the same for both stocks. Under this condition, δ is equal to 1 to adjust for their mechanical differences in share price.

[21] A stock whose price is below $1 has a tick size smaller than 1 cent and we find that 10 stocks have lot sizes of fewer than 100 shares because their prices are very high (e.g., Berkshire Hathaway A share).

and Quote (TAQ) data to compute time-weighted bid–ask spreads and trading volumes. We use Center for Research in Security Prices (CRSP) data to compute volatility and daily average prices in a given year. We winsorize our variables at the 1% level. Table 2 presents the summary statistics for our sample.

TABLE 2
Summary Statistics

	Mean	Min	Q1	Median	Q3	Max	Std.Dev	N
Nominal Price ($)	43.37	1.00	7.66	17.64	43.96	2884.50	100.72	3745
Spread (cents)	16.84	1.00	2.82	6.48	16.13	318.07	32.47	3745
Log(#Trades)	7.89	2.40	6.77	8.11	9.12	12.64	1.77	3615
Log(Volatility)	0.05	0.01	0.03	0.04	0.06	0.19	0.03	3745
Log(Dollar Volume)	22.19	12.26	20.25	22.32	24.11	30.14	2.64	3745
Log(Market Cap)	20.45	14.67	18.93	20.34	21.88	28.05	2.13	3745

In this table we report the summary statistics for our U.S.-listed stock sample for cross-sectional tests. We take a snapshot of the year 2020 as our sample, and we take the annual averages of the data. We require the stocks to have the standard 100-share lot size, a price above $1 over the course of the entire year, and at least 20 observations within the year.

We present our regression results in Table 3. The results reported in Panel A of Table 3 strongly reject the null hypothesis that $\delta = 1$. Therefore, the percentage bid–ask spread depends strongly on the dollar lot size. The results reported in Column (1) show that $\delta = 2.09$, which is close to our model's prediction of $\delta = 2$. The coefficients for volatility and trading volume are quantitatively close to 1. Our parsimonious three-variable model captures most of the cross-sectional variation in the bid–ask spread, with an R^2 as high as 0.81. Columns (2)–(5) show similar results using years prior to 2020, indicating the robustness of the Modified Square Rule.

We use Panel B of Table 3 to compare our three-variable liquidity model with two canonical benchmarks: Madhavan (2000) and Stoll (2000). These benchmarks include all three variables in our model,[22] plus market cap (Madhavan 2000 and Stoll 2000) and the

[22] Madhavan (2000) measures volatility using standard deviations as we do and Stoll (2000) measures volatility using variance.

number of trades (Stoll 2000). We compute market caps using CRSP and the number of trades using TAQ.

TABLE 3
Lot-driven Spread and the Modified Square Rule

Panel A: Three-Variable Model of Liquidity

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	$Log(\overline{s_t^L}) = Log(\overline{s_t^{tot}} - \Delta)$					
Sample Period	2020	2019	2018	2017	2016	2015
$Log(\overline{Price})$	**2.09*****	**2.08*****	**2.12*****	**2.06*****	**2.08*****	**2.06*****
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$Log(\overline{Volatility})$	**0.96*****	**1.19*****	**1.16*****	**1.07*****	**1.17*****	**1.14*****
	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$Log(\overline{DVol})$	**-0.84*****	**-0.82*****	**-0.81*****	**-0.79*****	**-0.83*****	**-0.81*****
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Obs.	3745	3652	3736	3711	3713	3850
R^2	0.8063	0.8389	0.8003	0.7704	0.8095	0.8298
Adj. R^2	0.8061	0.8387	0.8001	0.7702	0.8093	0.8296

Panel B: Specification Horseraces for the Three-Variable Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable	$Log(\overline{s_t^L})$	$\frac{\overline{s_t^{tot}}}{2}$(bps)	$\frac{\overline{s_t^{tot}}}{2}$(bps)	$Log(\overline{s_t^L})$	$Log(\overline{s_t^L})$	$Log(\overline{s_t^L})$	$Log(\overline{s_t^L})$
Sample Period	2020	2020	2020	2020	2020	2020	2020
$Log(\overline{Price})$	**2.09*****		**6.20*****	**2.24*****	**2.32*****	**2.26*****	**2.38*****
	(0.03)		(1.56)	(0.03)	(0.04)	(0.03)	(0.03)
$Log(\overline{Volatility})$	**0.96*****					**0.68*****	0.03
	(0.05)					(0.06)	(0.06)
$Log(\overline{DVol})$	**-0.84*****	**-30.46*****	**-19.17*****	**-0.57*****	**-0.75*****	**-0.60*****	
	(0.02)	(1.49)	(2.67)	(0.03)	(0.07)	(0.03)	
$Log(\overline{MKTCAP})$		**19.08*****	**9.69*****	**-0.46*****	**-0.42*****	**-0.41*****	**-1.22*****
		(1.71)	(1.21)	(0.04)	(0.03)	(0.04)	(0.02)
$Log(\overline{\#Trades})$			**-7.12*****		**0.10*****		
			(2.72)		(0.07)		
$\overline{Volatility} \times 10^2$		**5.40*****		**0.09*****			
		(0.53)		(0.01)			
$\overline{Variance} \times 10^4$			**0.19*****		**0.01*****		
			(0.02)		(0.00)		
$\overline{Price}^{-1}$		**-39.91*****					
		(5.17)					
Obs.	3745	3745	3745	3745	3745	3745	3745
R^2	0.8063	0.6191	0.6529	0.8133	0.8228	0.8207	0.7478
Adj. R^2	0.8061	0.6187	0.6524	0.8131	0.8226	0.8205	0.7476

In this table we report the results of testing the Modified Square Rule on the cross-section of U.S. common stocks. In Panel A, we report the results of regressing the log of time-weighted lot-driven nominal spreads on the log of nominal prices, controlling for log(Volatility) and log(Volume). We take annual snapshots of 2016-2020 of U.S.-listed common stocks as our sample, and we take the annual averages of the daily data. We require the stocks to have the standard 100-share lot size, a price above $1 over the course of the entire year, and at least 20 observations within the year. In Panel B, we use the results to compare the Modified Square Rule with alternative specifications. In Column (1), we report the results derived with our model, while for columns (2) and (3) we incorporate the specifications of Madhavan (2000) and Stoll (2000), respectively. For columns (4) and (5), we use our specification to control for the price while keeping all other specifications in Madhavan (2000) and Stoll (2000) the same. For columns (6)–(8), we estimate our model with alternative control variables. Statistically significant coefficient estimates are shown in bold and standard errors are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The results reported in column (1) show that the R^2 of the three-variable model (0.81) is much higher than that in Madhavan (2000; see column 2, 0.62) and Stoll (2000; see column 3, 0.65), even though our three-variable model includes only a subset of variables that have been included in previous benchmarks. The results displayed in columns (4)–(8) provide two explanations for this surprising outperformance. First, our three-variable model finds a better functional form to control for price. Second, our three-variable model removes redundant variables in the regression.

First, our model provides a better functional form to control for price. Madhavan (2000) uses $price^{-1}$ to control for the relative tick size, while Stoll (2000) uses $log(price)$. Both specifications impose a monotonic relationship between the price and the percentage spread. In reality, the relationship between price and liquidity is U-shaped (Mackintosh 2021). Therefore, these two canonical benchmarks may misspecify the relationship between price and liquidity, at least for recent years. One indicator of the misspecification is the coefficient estimate of the price. Both Madhavan's (2000) and Stoll's (2000) specifications show that an increase in the price increases the percentage spread. The positive correlation works exactly in the opposite direction of their economic reasoning. Madhavan (2000) and Stoll (2000) use price as a proxy for the relative tick size. An increase in price, or a decrease in the relative tick size, should not reduce liquidity. Indeed, overwhelming evidence shows that exogenous decreases in the tick size reduce the percentage spread (see Bessembinder 2003; Albuquerque, Song, and Yao 2020, for example.).

The economic factor that reconciles this contradiction is the dollar lot size. Exogenous changes in the tick size do not change the dollar lot size. Therefore, Bessembinder (2003) and Albuquerque, Song, and Yao (2020) show that an increased tick size increases the percentage spread while holding the dollar lot size fixed. A price increase, however, reduces the relative tick size but increases the dollar lot size.

Our model indicates that a better functional form in the regression would be to subtract one tick from the bid–ask spread to control for the tick size and use log (price) to control for the lot size. To obtain the results reported in columns (4) and (5), we use our specification to control for the price while keeping all other specifications in Madhavan (2000) and Stoll (2000) the same. The R^2 in Madhavan's (2000) specification increases from 0.62 to 0.81, while the R^2 in Stoll's (2000) specification increases from 0.65 to 0.82.

Second, our model enables us to remove redundant explanatory variables such as the market cap. Almost all empirical tests of liquidity control for the market cap, reflecting the intuition that large-cap stocks should be more liquid (Stoll 2000; Madhavan 2000). The results reported in column (6) show that adding the market cap to our three-variable model increases the R^2 by only 0.01. The results reported in column (7) show that the R^2 declines from 0.81 to 0.75 if we remove the dollar volume but keep the market cap. Our model provides the intuition that explains why the market cap has almost no additional explanatory power for the percentage spread. Notice that we model market cap as v_t, and it affects liquidity only through its product with λ_I, the turnover rate. Therefore, our model indicates that the market maker cares more about the dollar trading volume that pays the bid–ask spread and less about firm size *per se*. A small-cap stock with high turnover is as liquid as a large-cap stock with low turnover if they have the same dollar volume because the competitive market maker earns the same revenue and *ceteris paribus* quotes the same bid–ask spread.

Stoll's (2000) and Madhavan's (2000) specifications also report that an increase in the market cap increases the percentage spread, whereas economic intuition suggest the opposite. In summary, the weaker explanatory power of the market cap and its contradictory coefficient both indicate that the market cap is a redundant variable after controlling for the dollar volume. This result is consistent with our model prediction: although the market cap appears to be a universal explanatory variable in most regressions, it does not directly affect the market maker's decision regarding the bid–ask spread after controlling for the dollar volume.

7. A TWO-VARIABLE EMPIRICAL MODEL OF NOMINAL PRICES

In this section we show that our model captures 57% of the cross-sectional variation in stock prices. In addition to this considerable explanatory power, our results also address two puzzles in the behavioral finance literature.

Proposition 3 predicts that a firm's optimal nominal price is $p^* = \sqrt{\frac{\lambda_I v \Delta}{2\sigma\lambda_J L}}$. Taking the natural log on both sides, we obtain:

$$log(p^*) = \frac{1}{2}log(\lambda_I v) - \frac{1}{2}log(\sigma\lambda_J) + const. \tag{21}$$

We test Equation (21) using the daily average of each variable in the year 2020.

$$log(\overline{Price})_i = \alpha + \frac{1}{2}log\left(\overline{DVol}\right)_i - \frac{1}{2}log\left(\overline{Volatility}\right)_i + \varepsilon_i. \tag{22}$$

In Table 4 we report the test results for nominal prices. In column (1), our reported results show that the two-variable model of volatility and dollar volume captures 57% of the cross-sectional variations in stock prices.

TABLE 4
Two-Variable Model of Nominal Prices

	(1)	(2)	(3)	(4)
Dependent Variable		$Log(\overline{Price})$		
$Log(\overline{Volatility})$	**-0.99*** **			**-0.93*** **
	(0.04)			(0.04)
$Log(\overline{DVol})$	**0.29*** **	**0.33*** **	**0.34*** **	**0.30*** **
	(0.01)	(0.01)	(0.03)	(0.01)
Industry FE	N	N	Y	Y
Obs.	3745	3745	3745	3745
R^2	0.5664	0.4387	0.4811	0.5787
Adj. R^2	0.5661	0.4386	0.4775	0.5757

In this table we report the results of testing the two-variable model on the cross-sectional nominal prices of U.S. common stocks. We take a snapshot of U.S.-listed common stocks in the year 2020 as our sample and we take the annual average of the data for daily observations. We require the stocks to have the standard 100-share lot size, a price higher than $1 during the entire year, and at least 20 observations within the year. In Column (1), we report the results derived from our model. The results reported in Columns (2), (3), and (4) enable us to compare the predictive power of industry fixed-effects as suggested by Weld et al. (2009) with volatility. Statistically significant coefficient estimates are shown in bold and standard errors are shown in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Our model also rationalizes two puzzles in behavioral finance. One is the volatility puzzle raised by Baker, Greenwood, and Wurgler (2009), who hypothesize that volatile firms are less likely to split their stocks because they have a "*greater chance of reaching a low price anyway.*" Therefore, Baker, Greenwood, and Wurgler (2009) find that *"a somewhat unexpected result is the effect of volatility, which suggests that volatile firms have a greater, not lesser, propensity to manage prices downward."* We reconfirm their finding in our sample, and the result is consistent with the predictions generated by our model.

We explain this result using the Two-Tick Rule. A universal optimal nominal bid–ask spread implies that any fundamental variable that increases the percentage spread should reduce the nominal price. The nominal price decreases with volatility because an increase in volatility increases the market maker's adverse selection risk and thereby the percentage spread.

We can also interpret this result using lot constraints or tick constraints. Because an increase in volatility increases the adverse-selection risk for market makers, firms whose volatility is higher should choose a lower price to reduce the dollar lot size. Also, firms that experience greater volatility have wider percentage spreads. Therefore, they can choose lower prices because their bid–ask spreads are constrained to a lesser extent by the tick size.

The second puzzle is the social norm puzzle suggested in Weld et al. (2009). After ruling out several alternative hypotheses that may explain the nominal price, the authors propose that customs and norms can explain the nominal price puzzle. One piece of evidence of the effect of social norms is that firms usually choose prices that are consistent with those their industry and size peers choose.

Weld et al. (2009) use industry fixed effects to explain nominal prices in the cross-section. We reconfirm their results, as reported in columns (2) and (3) of Table 4. Starting with a univariate regression with log volume and adding industry fixed effects increases the R^2 from 0.44 to 0.48, so the marginal contribution of industry fixed effects is 0.04. When we add volatility to the regression, though, the industry fixed effects increase the R^2 by only 0.01, as reported in column (1) and column (4). Therefore, volatility subsumes most of the explanatory power of industry fixed effects. Therefore, a rational interpretation of the industry clustering is that firms that operate in the same industry may be subject to similar volatility.

Consistent with our model, we show in column (1) of Table 4 that the nominal price also increases with the dollar volume. As larger stocks tend to trade in higher volumes, the results reported in column (1) explain the stock-price clustering around size. Moreover, our model explains why larger stocks cluster at higher prices. Holding all else equal, increasing the dollar volume increases the market maker's revenues and thereby the percentage spread. A uniform optimal nominal bid–ask spread then implies that active stocks should choose higher nominal prices. Again, we can also interpret the positive relationship between the dollar volume and the price using tick- or lot-size constraints. For example, an increase in dollar volume increases the market maker's revenue and reduces the percentage spread.

When percentage spread falls, tick size becomes a more binding constraint. Therefore, firms should choose higher prices to relieve tick-size constraints.

In summary, our model fits qualitatively with cross-sectional variations in nominal prices. The fit is not as good as the fit for the bid–ask spread (0.57 vs. 0.81), and the coefficient on the estimate does not change one for one with model predictions. Interestingly, this imperfect fit enables us to identify the impact of prices on the bid–ask spread. If all firms chose their prices following our model, $log(\overline{price})$ would correlate almost perfectly with $log(\overline{Volatility})$ and $log\ (\overline{DVol})$, leading to collinearity.

There are two possible, albeit not mutually exclusive, interpretations to explain why the goodness of fit of firms' behavior is less perfect than that of traders' behavior. First, we model the price that optimizes liquidity, whereas a firm may have other objectives. For example, Berkshire Hathaway A shares *"try to avoid policies that attract buyers with a short-term focus on our stock price"* (Buffett 1983). A higher price and a higher transaction cost both help for this purpose. Indeed, we find that Berkshire Hathaway's A shares have a very high transaction cost of 9.73 bps, which is 5.2 times higher than that of its B shares. Using the Modified Square Rule, we find that the difference in their tick and lot sizes can almost fully explain their 5.2-time difference in transaction costs.[23] More interestingly, we find that the nominal bid–ask spread of BRK.B becomes 1.46 cents in the year after Berkshire Hathaway split the stock in 2010, which is close to what the Two-Tick Rule would do. Therefore, although Berkshire Hathaway may use the high share price for its A shares to serve other purposes, it did choose a liquidity-maximizing price for its B shares.

Second, firms respond less dramatically to market-structure frictions than traders do. Therefore, firms may end up with suboptimal nominal prices. We cannot separate this hypothesis from the previous one, but the benchmark of the liquidity-maximizing price is important in both cases. If a firm does not have any other objective than to manage its price,

[23] Class A shares of Berkshire Hathaway (BRK.A) have a lot size of 1 share, whereas Class B (BRK.B) have a normal lot size of 100 shares. The BRK.A traded at a \$112.10 bid–ask spread in 2011, the year after BRK.B's split, and each Class A share represents 1,500 Class B shares. If BRK.A splits 1500-for-1 and then increases its lot size by 100 times, its predicted lot-driven spread would decrease to 0.50 cents ($\frac{11209}{1500^2} \times \frac{100}{1}$). By adding the one cent tick-driven spread, the total spread becomes 1.50 cents, which is close to BRK.B's observed spread as well as the optimal two-cent spread.

it may manage the price to approach the price that maximizes liquidity. If a firm chooses another price because of other objectives, our paper provides an estimation of the liquidity cost to achieve other objectives.

8. STOCK SPLITS AND SPLIT ANNOUNCEMENT RETURNS

Firms can adjust their stock prices using stock splits. In this section, we show that our model rationalizes stock splits and split-announcement returns. In Subsection 8.1, we describe our data and sample. In Subsection 8.2, we find that moving to the optimal price predicted by our model can rationalize more than 90% of stock splits. We also show that changes in the percentage spreads after splits match up almost one for one with our model predictions. In Subsection 8.3, we find that model-predicted changes in percentage spreads can explain 94 bps, or more than one-third, of split-announcement returns. Finally, our back-of-the-envelope calculation in Section 8.4 shows that the median U.S. stock value would increase by 106 bps and the total U.S. market capitalization would increase by $93.7 billion if all firms split to their optimal prices.

8.1 Data, Sample, and Summary Statistics

Our sample includes all U.S. common stock–split announcements (CRSP event code 5523) from June 2003 through December 2020.[24] We exclude reverse splits, for two reasons. First, CRSP does not record reverse-split announcement dates. Second, reverse splits are usually mechanical and associated with bad news. For example, one major reason for reverse splits is that firms must comply with the minimum listing requirement of a $1.00 minimum bid price (Martell and Webb 2008).

We require stocks to be U.S.-listed common stocks (the SHRCD is 10 or 11) and have pre- and post-split prices higher than $1 per share. We use CRSP data for stock-split ratios, split-announcement dates, split-adjusted stock returns, market returns around declaration dates, and other control variables. We use millisecond TAQ data to calculate time-weighted

[24] The sample period begins in the month when the millisecond TAQ data become available.

quoted bid–ask spreads. To calculate cumulative abnormal returns (CARs), we obtain daily Fama-French factor returns and risk-free rates from Kenneth French's data library. We also require that the declaration date, the ex date, and the split ratio be neither missing nor duplicated by CRSP. In addition, we use COMPUSTAT data to obtain annual reported numbers of shareholders, and we aggregate 13-F filings to calculate the institutional holdings of a stock one quarter before and after its split announcement. Variables are winsorized at the 1% level. Following Grinblatt, Masulis, and Titman (1984), we require that stock-split ratios be greater than or equal to 1.25 (5-for-4). We end up with 1,196 stock splits.

In Table 5 we report the descriptive statistics. Our sample comprises 912 individual stocks. The most common splits are 2-for-1 splits (649 times) and 1.5-for-1 splits (359 times), and the mean split ratio is 1.91. The average price before a split announcement is $59.49, and the average price after a split is $33.15. The number of trades increases by 74%, but the dollar trading volume almost does not change. This result supports our model. The dollar trading volume does not change much because it reflects traders' intrinsic needs, but the number of trades increases post stock splits because similar demand is sliced into smaller and more frequent units. We find that institutional holdings increased slightly, from 57.90% to 58.04%, indicating that retail traders' holdings do not change dramatically. Therefore, changes in the compositions of retail/institutional holdings are unlikely to drive our results. We find that split announcements lead to an average abnormal return of 2.73%, whereas the average abnormal return around the ex dates is only 0.30% and the median abnormal return around ex dates is only -0.01%. The results are consistent with Fama et al. (1969), who find that new information contained in stock splits is incorporated into the price immediately after its announcement. Table 5 also provides preliminary evidence that splits move the bid–ask spread toward the two-tick optimum. The average bid–ask spread falls from 15.69 cents to 9.06 cents and the median bid–ask spread falls from 7.02 cents to 4.30 cents.

TABLE 5
Summary Statistics

	Mean	Min	Q1	Median	Q3	Max	Std.Dev	N
Split Ratio	1.91	1.25	1.50	2.00	2.00	50.00	1.52	1196
Pre-split price ($)	59.49	3.49	31.79	47.58	70.75	3311.00	105.02	1196
Post-split price ($)	33.15	2.15	21.45	30.01	40.30	440.64	20.28	1196
Market cap ($bn)	8.54	0.01	0.42	1.51	4.44	2206.91	68.60	1196
Ex-ante spread (cents)	15.69	1.04	4.20	7.02	16.60	294.54	24.77	1196
Ex-post spread (cents)	9.06	1.03	2.79	4.30	8.98	103.22	12.44	1196
Announcement CAR (%)	2.73	-28.7	-0.13	1.79	4.25	68.83	5.88	1196
Ex-date CAR (%)	0.30	-38.9	-1.90	-0.01	2.01	156.68	6.50	1196
Ex-ante volume ($MM)	39.69	0.01	1.19	9.12	35.28	491.63	80.15	1196
Ex-post volume ($MM)	41.45	0.01	1.62	10.42	38.61	486.61	82.75	1196
Pre-split trades (thousands)	3.39	0.00	0.29	1.03	2.77	365.77	15.70	1141
Post-split trades (thousands)	5.93	0.00	0.48	1.56	4.32	852.72	34.72	1163
Pre-split inst. holding (%)	57.90	0.00	32.18	66.05	84.44	99.11	30.98	942
Post-split inst. holding (%)	58.05	0.00	32.89	66.06	84.19	97.83	30.13	924
Pre-split holders (thousands)	14.91	0.00	0.45	1.95	7.18	1234.00	65.86	753
Post-split holders (thousands)	16.45	0.00	0.44	2.04	7.80	1426.00	71.60	758
Log(holders change ratio)	0.03	-6.93	-0.10	-0.01	0.20	4.08	0.80	729

In this table we report the summary statistics for our stock-split sample for September 2003–December 2020. Institutional holdings are taken from 13-F filings for the quarters immediately before and after split-announcement dates. Following Grinblatt, Masulis, and Titman (1984), announcement and ex-date CARs are cumulated announcement returns during dates [T-1, T+1]. Shareholder numbers are taken from the years immediately before and after stock-split announcements, and, following Amihud, Mendelson, and Uno (1999) the logs of the changes are reported. Other pre-split variables are measured in 180-day-to-60-day windows before split-announcement days and post-split variables are measured in 60-day-to-180-day windows after split-implementation days.

8.2. Explaining Stock Splits

The Modified Square Rule (Corollary 2) indicates that the change in the percentage spread after splits will be

$$R_i = \frac{(s_i^{pre} - \Delta)/H_i^2 + \Delta}{p_i^{pre}/H_i} - \frac{s_i^{pre}}{p_i^{pre}}, \tag{23}$$

where $\frac{s_i^{pre}}{p_i^{pre}}$ is the percentage spread before splits and $\frac{(s_i^{pre} - \Delta)/H_i^2 + \Delta}{p_i^{pre}/H_i}$ is the post-split percentage spread predicted by the Modified Square Rule. Therefore, our model predicts that a split is correct if $R_i < 0$ and a split is incorrect if $R_i > 0$. We find that 1,089 splits are "correct" and 107 splits are "incorrect." Therefore, the move toward liquidity-optimizing prices explains at least 91% of stock splits ($\frac{1089}{1089+107}$). Among the 107 incorrect splits, 74 should have split because their bid–ask spreads are wider than the two-tick optimum. They choose split ratios that are so aggressive, though, that their new bid–ask spreads are further away from the two-tick optimal. We find that R_i, on average, decreases by 15.22 bps in our sample, providing additional evidence that one possible goal of stock splits is to improve liquidity.

When firms make decisions on stock splits, they do not know the realized liquidity change post splits. Yet we can observe the realized change in percentage spreads ex post. Whether predicted changes match realized changes provides a test for our model. We define the realized change in the percentage spread, ΔS_i, as the difference between the average percentage spread 180 to 60 days before announcement days and the average percentage spread 60 to 180 days after ex dates. [25] Following Weld et al. (2009), our control variables include market capitalization, price, volume, and turnover rates.

The results we report in Table 6 show that a predicted one bp increase in the spread leads to a 0.94 bps realized increase (column 1), with a *t*-statistic of 4.28. Therefore, the Modified Square Rule strongly predicts the percentage spread after splits. The signaling literature hypothesizes that a more aggressive split ratio should reduce liquidity to a greater extent and send a stronger signal to the market (Mcnichols and Dravid 1990). The results reported in columns (2) and (3) indicate, however, that the split ratio does not predict the change in liquidity in our sample. The results reported in column (3) indicate that a one bp increase in the predicted spread leads to a 1.02 bps realized increase in the percentage spread after controlling for the split ratio.

[25] Stock trades around split announcements are volatile (Ohlson and Penman, 1985). Therefore, when measuring the bid–ask spread, we exclude 60 days around the split window and consider the spread difference between the two relatively calm periods before the announcement and after the ex date (the day that the split actually happens).

TABLE 6

Predictions of Changes in Bid–Ask Spreads

Dependent Variable	Realized ΔS_i (bps)		
	(1)	(2)	(3)
R_i *(bps)*	**0.94*****		**1.02*****
	(0.22)		(0.20)
$Log(H_i)$		-0.10	0.17*
		(0.07)	(0.10)
Controls	Y	Y	Y
Industry-Year FE	Y	Y	Y
Obs.	1196	1196	1196
Adj. R^2	0.331	0.198	0.336

In this table we report the results obtained from regressing realized changes in the percentage spread on predicted spread changes. R_i is the model-predicted change in the percentage spread (in bps). We control for the split ratio, which comes from CRSP item FACSHR. Following Weld et al. (2009), other control variables include log(market cap), price, log(volume), and turnover rates. We also control for industry-year fixed effects to absorb any industry-year-specific shocks, where each industry is defined by reference to the first two digits of the NAICS classification. Statistically significant coefficient estimates are shown in bold and standard errors are shown in parentheses. Standard errors are adjusted for both heteroskedasticity and within correlations clustered by firm. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

8.3 Explaining Split Announcement Returns

As liquidity affects asset values (Amihud and Mendelson, 1986), the liquidity improvement predicted by our model contribute to the positive announcement returns of stock splits. Although the median time between the announcement and ex dates is 24 business days, in the previous subsection we show that the Modified Square Rule can predict the realized liquidity change almost one for one. Therefore, our results explain why the abnormal return is realized immediately after the announcement.

Figure 2 presents preliminary evidence that our model-predicted liquidity change affects returns on split announcements. Firms in the group with predicted correct splits realize an average announcement CAR of 2.87%, whereas those in the group with predicted incorrect splits obtain an average announcement return of only 1.36%.[26]

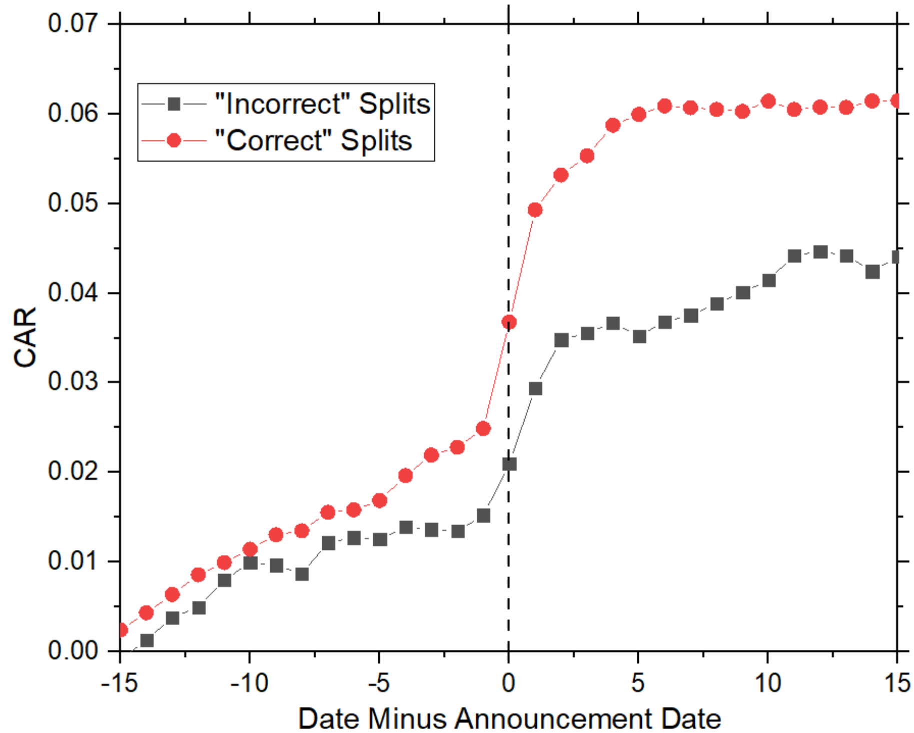

FIGURE 2.—Split-Announcement Returns: In this figure we report the cumulative abnormal returns (CARs) around split-announcement dates. Our sample includes all U.S.-listed common stock splits from September 2003 through December 2020. We require a firm to choose at least a \$1 nominal price before and after a split. We categorize stocks into two types based on the prediction from Equation (23). A split is "correct" if the predicted percentage spread decreases and "incorrect" if the predicted percentage spread increases.

Insofar as splits are good news in general (Fama et al. 1969, Brennan and Copeland 1988, Lamoureux and Poon 1987, Maloney and Mulherin 1992), both groups enjoy positive returns, but the 1.51% difference indicates that predicted liquidity changes may contribute to the difference in returns. To test this hypothesis, we run the following regression:

$$CAR_{i,[T-1,T+1]} = \theta \cdot R_i + Controls_i + Industry_i \times Year\ FE_t + \varepsilon_i. \qquad (24)$$

Again, to reflect the information sets of traders on stock-split-announcement days and to avoid look-ahead bias, we use the predicted spread change R_i but not the realized spread

[26] Following Grinblatt, Masulis, and Titman (1984), we consider the window of announcement abnormal returns as dates -1, 0, and 1.

change. Following Weld et al. (2009), our control variables include market capitalization, price, volume, and turnover rates. As an additional robustness check, we also control for industry-year fixed effects to absorb any industry and time-specific shocks, where each industry is defined by reference to the first two digits of the NAICS classification. We also control for institutional holding changes and the number of investor changes (following Amihud, Mendelson, and Uno (1999) and Dyl and Elliott (2006)) to control for the impact of the investor base.[27]

The results reported in Table 7 show that our predicted spread change is significantly negatively associated with split-announcement abnormal returns. The results reported in column (1) indicate that a predicted one bps increase in the percentage spread is associated with -5.47 bps in announcement returns.[28] After adding control variables, the results reported in column (5) indicate that a predicted one bps increase in the percentage spread is associated with -6.18 bps in announcement returns. As the mean of R_i is -15.22 bps, correct split ratios contribute $-15.22 \times -6.18 = 94$ bps to the overall average split-announcement abnormal return of 273 bps. Therefore, a reduction in market-microstructure friction partially explains why a seemingly cosmetic change, a stock split, leads to positive returns.

The Table 7 results show that the explanatory power of the tick-and-lot channel is orthogonal to two existing interpretations of splits and announcement returns from splits. Brennan and Copeland (1988) propose that firms use splits to convey positive signals about firm fundamentals, and the cost of such signals is reduced liquidity. The signaling channel predicts that a larger reduction in liquidity should send a stronger signal and be associated with higher returns. We find, however, that splits improve liquidity and the column (1) results indicate that a greater improvement in liquidity leads to a higher return. The signaling channel also predicts that a larger split ratio should send a stronger signal and be

[27] These two variables are missing for more than half of the firms, so we do not add them in the baseline test. As the results reported in column (4) of Table 7 show, our results are robust to these additional controls in the reduced sample.

[28] The economic magnitude is similar to that reported in Albuquerque, Song, and Yao (2020). Using a controlled experiment, they find that a 43.5 to 48.2 bps increase in the bid–ask spread led to a 175 to 320 bps drop in asset values.

associated with higher returns (Mcnichols and Dravid 1990). As reported in column (2), we find that an increase in the split ratio leads to significantly higher returns, but the column (3)–(6) results indicate that the predictive power disappears after controlling for R_i.

Lamoureux and Poon (1987) and Maloney and Mulherin (1992) propose that firms use stock splits to attract retail traders, and an increase in uninformed traders increases volume and liquidity. As seen in Table 5, we find that institutional holdings increase slightly after stock splits. The results reported in column (6) of Table 7 show that the change in retail holdings, proxied by the number of shareholders and institutional holdings, does not affect announcement returns.[29]

TABLE 7
Predicted Spread Changes and Abnormal Returns on Announcements

Dependent Variable	$CAR_{i,[T-1,\ T+1]}$ (bps)					
	(1)	(2)	(3)	(4)	(5)	(6)
R_i *(bps)*	**-5.47*****		**-5.42*****	**-4.42****	**-6.18*****	**-5.90****
	(1.48)		(1.47)	(2.06)	(2.40)	(2.87)
$Log(H_i)$		**1.40****	1.12*	1.93	1.46	2.00
		(0.67)	(0.63)	(1.29)	(1.15)	(1.41)
$Log(MKTCAP_i)$				**4.52*****	**4.81****	**9.43*****
				(1.59)	(1.80)	(2.16)
$Log(Price_i)$				**-5.80*****	**-6.29*****	**-10.72*****
				(1.63)	(1.76)	(2.14)
$Turnover_i$				**5.53*****	**5.63*****	**10.03*****
				(1.59)	(1.81)	(2.12)
$Log(Volume_i)$				**-4.87*****	**-4.96*****	**-9.51*****
				(1.54)	(1.72)	(2.06)
$Log\left(\frac{InstHldg_{After}}{InstHldg_{Before}}\right)$						5.43*
						(3.19)
$Log\left(\frac{TOTSH_{After}}{TOTSH_{Before}}\right)$						-0.18
						(0.26)
Industry-Year FE	N	N	N	N	Y	Y
Obs.	1196	1196	1196	1196	1196	607
Adj. R^2	0.067	0.003	0.070	0.132	0.164	0.262

[29] Column (6) reflects only half of the observations because of missing observations for institutional holdings in the 13-F and missing observations for number of shareholders items in COMPUSTAT. To account for missing observations, we use the sensitivity of R_i for column (5) in our later estimations.

In this table we report the results of regressing split-announcement CARs on predicted spread changes and announced split ratios with various controls. R is the model-predicted change in the percentage spread. The split ratio comes from the CRSP item FACSHR. Following Weld et al. (2009), for column (4) we control for log(market cap), price, log(share volume), and turnover rates before the splits. Industry-year fixed effects are added to obtain the results reported in Column (5) to absorb any industry-year-specific shocks, where each industry is defined by reference to the first two digits of the NAICS classification. For column (6), following Dyl and Elliott (2006) and Amihud, Mendelson, and Uno (1999), we also control for changes in institutional holdings and the numbers of shareholders. Institutional holdings are aggregated from quarterly 13-F filings before and after split announcements, and the numbers of shareholders are obtained from the COMPUSTAT annual item CSHR. Statistically significant coefficient estimates are shown in bold and standard errors are shown in parentheses. Standard errors are adjusted for both heteroskedasticity and within correlations clustered by firm. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

8.4. Back-of-the-Envelope Calculation

After showing that the percentage spread changes predicted by the Modified Square Rule match up almost one for one with the realized percentage spread changes, we carry out a back-of-the-envelope calculation to estimate the optimal price for each firm and the benefit gained by moving to the optimal price. The beauty of the Modified Square Rule and the Two-Tick Rule is that we do not need to calibrate firm fundamentals to solve for the optimal price. The only input we use is the time-weighted average nominal spread $\mathbb{E}(s_t^{tot})$, and our liquidity-optimal split ratio is $H^* = \sqrt{\frac{\mathbb{E}(s_t^{tot})-\Delta}{\Delta}}$. We use the ratio to adjust the stock's current price to obtain its optimal price. We find that the median optimal price for large stocks (the top tercile with NYSE breakpoints) is \$37.31, whereas the median small stock (the bottom tercile with NYSE breakpoints) can sustain an optimal price of only \$3.51.

We then insert the H^* in formula (23) to calculate the expected liquidity improvement from adjusting to the optimal price. After adopting optimal pricing, we predict that the median spreads will be reduced from 42.85 bps to 25.66 bps, a 40% reduction. Because the sensitivity of prices to liquidity changes that we report in Table 7 is 6.18, we expect the value of the median U.S. stock to increase by (42.85-25.66) × 6.18 = 106 bps after adopting optimal pricing. Summing up the potential gains for each stock, the total benefit of adopting optimal pricing is estimated to be $93.7 billion.

Not surprisingly, large stocks are the biggest winners in dollar amounts because they are actively traded. In our version dated before June 3, 2022, the top two winners are Amazon and Alphabet Inc. Consistent with our prediction, Amazon and Alphabet Inc. both announced major 20-for-1 stock splits. After these splits, the top winner would be UnitedHealth Group ($546 million if it splits from $302 to $80).[30] The adjustment to the optimal price will reduce the bid–ask spread for median large-cap stocks from 14.00 bps to 7.91 bps and increase their market value by 37.6 bps.

Surprisingly, the adjustment to the optimal price would lead to a ten times greater benefit in percentages for small-cap stocks than for large-cap stocks. Perhaps small-cap stocks can afford a much lower nominal price ($3.51). However, many small firms choose much higher prices, probably because people often consider a high-priced stock to be prestigious (Weld et al. 2009). Therefore, we find that medium-small stocks should adjust their price from $9.80 to $3.51. Their median spread would then decrease from 153 bps to 91.7 bps, and corporate value would increase by 378 bps. Certainly, the benefit in dollars is much smaller than that enjoyed by large-cap stocks. In fact, the smaller dollar gains may explain why firms in the treatment group of the Tick Size Pilot did not reverse split before the SEC increased their tick size to five cents and then split their stocks again after the pilot ended.[31]

---

[30] Interestingly, UnitedHealth Group may want to maintain its high nominal price because it is currently the largest component of the Dow Jones Industrial Average Index. The index uses simple average of stock prices, so a stock split will lower United Health's weight within the index. Again, our liquidity channel should not be considered the only determinant of nominal prices. Rather, our model quantifies the costs of deviating from the liquidity-maximizing prices.

[31] In 2016, the SEC selected 1,200 small and median stocks and increased their tick size from one cent to five cents for two years. Therefore, the tick-driven half spread increased by two cents. As the treatment group had a median trading volume of 16.3 million shares in 2017, the annual increase in transaction costs is capped at $326,000. The cost reduction from the first reverse split and then the split back is even smaller, and Corollary 4 shows that reverse splits can only partially reverse the negative impact of an increase in tick size. Weld et al. (2009) estimate that the fixed cost of a split ranges from $250,000 to $800,000. Therefore, the fixed costs might help explain why firms did not reverse split and then split back for the two-year pilot. On the other hand, we do find that stocks in the control group were 44% more likely to split than stocks in the treatment

9. CONCLUSION

Economic models often incorporate an implicit but important assumption—continuous pricing and continuous quantities. In this paper, we offer the first study where both prices and quantities are discrete, and we show that these two seemingly small frictions help to address questions and puzzles in market microstructure, behavioral finance, and corporate finance.

Regarding market microstructure, the Modified Square Rule explains 81% of cross-sectional variation in the bid–ask spread. Our three-variable model of liquidity outperforms two benchmarks (Madhavan 2000 and Stoll 2000) even though it uses a subset of their parameters. The key to this surprising outperformance comes from the functional form. Madhavan (2000) and Stoll (2000) assume a monotonic relationship between price and liquidity, following the intuition that an increase in price reduces the relative tick size. Our theoretical model and empirical specification capture the U-shaped relationships between price and liquidity by discovering the trade-off between discrete pricing and discrete quantities. We encourage researchers to consider our empirical specification when they search for new explanatory variables that are related to liquidity: using the bid–ask spread minus one tick as the dependent variable to control for friction from tick-driven spreads and then using log(price) as the independent variable to control for frictions driven by the lot size.

Our paper explains 57% of cross-sectional variations in prices and resolves two puzzles in behavioral finance. Baker, Greenwood, and Wurgler (2009) find it unexpected that volatile stocks have a greater propensity to manage their prices downward. Our Two-Tick Rule provides an explanation for the volatility puzzle. As friction generated by a discrete price is always one tick, all firms reach the liquidity-maximizing price when their bid–ask spreads are two ticks wide, or when the friction from the tick size equals the friction from the lot size. As an increase in volatility increases the percentage spread,

group (13 vs 9). This comparison provides anecdotal evidence that an increase in tick size discourages stock splits.

the uniform two-tick optimal bid–ask spread implies a lower optimal price. We can also explain the volatility puzzle by reference to lot- or tick-size constraints. An increase in volatility increases the adverse-selection risk for market makers, and a firm should reduce its share price to relieve the lot-size constraints to reduce adverse-selection risk. Alternatively, an increase in volatility relieves tick-size constraints on prices and bid–ask spreads, giving firms more room to choose lower prices. Our paper also rationalizes why firms with similar market caps and firms that operate in the same industry choose similar prices (Weld et al. 2009). *Ceteris paribus,* larger firms trade in higher dollar volumes with lower percentage spreads, and they should choose higher prices to relieve tick-size constraints. We find that volatility largely subsumes industry fixed effects. Therefore, firms that operate in the same industry may choose similar prices because they experience similar volatility.

Regarding corporate decisions, our paper proposes a two-tick optimal rule for stock splits. We find that most stock splits move bid–ask spreads closer to two ticks. Therefore, our paper rationalizes stock splits. As an increase in liquidity increases stock value, we find that our tick-and-lot channel contributes 94 bps points to the average split-announcement return of 273 bps. Therefore, our paper also explains more than one-third of split-announcement returns. We estimate that the median value of a U.S. stock would increase by 106 bps if all firms were to move to their optimal prices, and total market value would increase by $93.7 billion. Finally, we find that realized changes in liquidity on ex dates match up almost one for one with the model predictions at announcement dates. This precise prediction rationalizes why most post-split returns occur at the announcement dates.

Our paper offers two policy implications. First, we discourage regulators from advancing the initiative to increase the tick size because it reduces liquidity, and we encourage them to advance the initiative to decrease the lot size because it improves liquidity. Second, we find that the move to a proportional tick-and-lot system reduces liquidity if regulators choose the tick and lot size for any existing stock under the uniform system as the benchmark. The economic intuition behind this surprising result is that the uniform system is actually more flexible than the proportional system. A uniform system allows firms to balance discrete prices with discrete quantities. A proportional system may reduce liquidity because it imposes the same level of discreteness on prices and quantities for all firms.

BRANDEIS UNIVERSITY
CORNELL UNIVERSITY

References

Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies*, 1(1), 3-40.

Albuquerque, R., Song, S., & Yao, C. (2020). The price effects of liquidity shocks: a study of the SEC's tick size experiment. *Journal of Financial Economics*, 138(3), 700-724.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid–ask spread. *Journal of Financial Economics*, 17(2), 223-249.

Amihud, Y., Mendelson, H., & Uno, J. (1999). Number of shareholders and stock prices: Evidence from Japan. *The Journal of Finance*, 54(3), 1169-1184.

Anshuman, V. R., & Kalay, A. (1998). Market making with discrete prices. The *Review of Financial Studies*, 11(1), 81-109.

Angel, J. J. (1997). Tick size, share prices, and stock splits. *The Journal of Finance*, 52(2), 655-681.

Asquith, P., Healy, P., & Palepu, K. (1989). Earnings and stock splits. *Accounting Review*, 387-403.

Back, K., & Baruch, S. (2004). Information in securities markets: Kyle meets Glosten and Milgrom. *Econometrica*, 72(2), 433-465.

Baker, H. K., & Gallagher, P. L. (1980). Management's view of stock splits. *Financial Management*, 73-77.

Baker, M., Greenwood, R., & Wurgler, J. (2009). Catering through nominal share prices. *The Journal of Finance*, 64(6), 2559-2590.

Baldauf, M., & Mollner, J. (2020). High-Frequency Trading and Market Performance. *The Journal of Finance*, 75(3), 1495-1526.

Bessembinder, H. (2003). Trade execution costs and market quality after decimalization. *Journal of Financial and Quantitative Analysis*, 747-777.

Blackrock (2019). Comment letter to the Securities Information Processor ("SIP") Operating Committees for the Unlisted Trading Privileges and Consolidated Quotation Plans. Available at: https://www.theice.com/publicdocs/BlackRock_Odd_Lot_Proposal_December_3_2019.pdf

Brennan, M. J., & Copeland, T. E. (1988). Stock splits, stock prices, and transaction costs. *Journal of Financial Economics*, 22(1), 83-101.

Brogaard, J., Hagströmer, B., Nordén, L., & Riordan, R. (2015). Trading fast and slow: Colocation and liquidity. *The Review of Financial Studies*, 28(12), 3407-3443.

Budish, E., Cramton, P., & Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics*, 130(4), 1547-1621.

Budish, E., Cramton, P., Kyle, A. and Lee, J. (2022). Flow Trading. Available at SSRN: https://ssrn.com/abstract=4145013.

Dyl, E. A., & Elliott, W. B. (2006). The share price puzzle. *The Journal of Business*, 79(4), 2045-2066.

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International Rconomic Review*, 10(1), 1-21.

Gensler,G. (2022). "Market Structure and the Retail Investor:" Remarks Before the Piper Sandler Global Exchange Conference. Available at https://www.sec.gov/news/speech/gensler-remarks-piper-sandler-global-exchange-conference-060822.

Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), 71-100.

Grinblatt, M. S., Masulis, R. W., & Titman, S. (1984). The valuation effects of stock splits and stock dividends. *Journal of Financial Economics*, 13(4), 461-490.

Kyle, A. S., Lee, J., (2017) Toward a fully continuous exchange, *Oxford Review of Economic Policy*, 33 (4), Winter 2017, 650–675.

Lakonishok, J., & Lev, B. (1987). Stock splits and stock dividends: Why, who, and when. *The Journal of Finance*, 42(4), 913-932.

Lamoureux, C. G., & Poon, P. (1987). The market reaction to stock splits. *The Journal of Finance*, 42(5), 1347-1370.

Li, S., Wang, X., & Ye, M. (2021). Who provides liquidity, and when?. *Journal of Financial Economics*, *141*(3), 968-980.

Mackintosh, P. (2022). Optimizing Trading via Market Structure. Available at https://www.nasdaq.com/articles/optimizing-trading-via-market-structure.

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets*, 3(3), 205-258.

Maloney, M. T., & Mulherin, J. H. (1992). The effects of splitting on the ex: A microstructure reconciliation. *Financial Management*, 44-59.

Martell, T. F., & Webb, G. P. (2008). The performance of stocks that are reverse split. *Review of Quantitative Finance and Accounting*, 30(3), 253-279.

McNichols, M., & Dravid, A. (1990). Stock dividends, stock splits, and signaling. *The Journal of Finance*, *45*(3), 857-879.

NASDAQ (2019). Comment letter to the Securities Information Processor (“SIP”) Operating Committees for the Unlisted Trading Privileges and Consolidated Quotation Plans. Available at: https://www.theice.com/publicdocs/SIP_Comment_Nasdaq_redacted.pdf

O'Hara, M., Yao, C., & Ye, M. (2014). What's not there: Odd lots and market data. *The Journal of Finance*, 69(5), 2199-2236.

Ohlson, J. A., & Penman, S. H. (1985). Volatility increases subsequent to stock splits: An empirical aberration. *Journal of Financial Economics*, 14(2), 251-266.

Shue, K., & Townsend, R. R. (2018). Can the Market Multiply and Divide? Non-Proportional Thinking in Financial Markets. *The Journal of Finance*.

Stoll, H. R. (2000). Presidential address: Friction. *The Journal of Finance*, 55(4), 1479-1514.

Yao, C., & Ye, M. (2018). Why trading speed matters: A tale of queue rationing under price controls. *The Review of Financial Studies*, *31*(6), 2157-2183.

Vayanos, D. (1999). Strategic trading and welfare in a dynamic market. *The Review of Economic Studies*, *66*(2), 219-254.

Weld, W. C., Michaely, R., Thaler, R. H., & Benartzi, S. (2009). The nominal share price puzzle. *Journal of Economic Perspectives*, 23(2), 121-42.

APPENDIX A: PROOFS

A.1. *Proof of Proposition 1 and Corollary 1*

The uninformed investors choose the arrival rate λ_q for all $q \in N^+$ subject to the total liquidity demand constraint $\sum_{q=1}^{\infty} qL\lambda_q = \lambda_I h$, where q is the order size in round lots. This proof shows that λ_q equals 0 for any $q \geq 2$ in equilibrium. That is, uninformed traders choose to slice their orders to minimum lots to achieve the lowest transaction costs. We first solve the market maker's quoted spreads for a given λ_q and then solve the uninformed trader's optimal λ_q through backward induction.

Observing the λ_I, λ_J, h, σ, and λ_q for $q \in N^+$, the competitive market maker quotes a liquidity schedule on both the bid and ask sides. We denote the spread for the q^{th} lot as s_t^q, and the market maker quotes at $p_t \pm \frac{s_t^q}{2}$. The q^{th} lot can trade only with liquidity-demanding orders that are larger or equal to q lots, which arrives at Poisson intensity $\sum_{i=q}^{\infty} \lambda_i$. As all market-maker quotes are subject to adverse selection risks with Poisson intensity λ_J, the probability of earning a spread of $\frac{s_t^q}{2}$ is $\frac{\sum_{i=q}^{\infty}\lambda_i}{\sum_{i=q}^{\infty}\lambda_i+\lambda_J}$ and the probability of being adversely selected is $\frac{\lambda_J}{\sum_{i=q}^{\infty}\lambda_i+\lambda_J}$. The equilibrium $\frac{s_t^q}{2}$ should equalize the revenue from providing liquidity and the loss from adverse selection, which leads to

$$\frac{\sum_{i=q}^{\infty}\lambda_i}{\sum_{i=q}^{\infty}\lambda_i+\lambda_J}\frac{s_t^q}{2} = \frac{\lambda_J}{\sum_{i=q}^{\infty}\lambda_i+\lambda_J}\left(\sigma p_t - \frac{s_t^q}{2}\right).$$

$$s_t^q = \frac{2\sigma p_t \lambda_J}{\sum_{i=q}^{\infty}\lambda_i+\lambda_J}. \qquad \text{for } q \in \mathrm{N}^+. \text{ (A.1)}$$

(A.1) implies that $s_t^1 \leq s_t^2 \leq \cdots \leq s_t^q \leq \cdots \leq 2\sigma p_t$, where the q^{th} equality holds only when $\lambda_q = 0$.

The transaction cost for uninformed traders is $\sum_{q=1}^{\infty}[C_B(q)+C_S(q)]\frac{\lambda_q}{2}$ per unit of time, where $C_B(q) = C_S(q) = \sum_{i=1}^{q}\frac{s_t^i}{2}$. Thus,

$$\sum_{q=1}^{\infty}[C_B(q)+C_S(q)]\frac{\lambda_q}{2} = \sum_{q=1}^{\infty}\frac{\lambda_q}{2}\cdot\sum_{i=1}^{q}s_t^i = \sum_{i=1}^{\infty}s_t^i\cdot\sum_{q=i}^{\infty}\frac{\lambda_q}{2}.$$

The second equality holds by swapping the order of summation. In other words, the spread s_t^i are paid at intensity $\frac{\lambda_i+\lambda_{i+1}+\cdots}{2}$. We show that, to minimize the transaction costs, the uninformed traders should choose only one order size q^*, so $\lambda_i = 0$ for all $i \neq q^*$. Otherwise, suppose there exist $q_1 < q_2$ such that $\lambda_{q_1} > 0$ and $\lambda_{q_2} > 0$. We have $s_t^{q_1} < s_t^{q_2}$ because the market maker quotes a wider spread for q_2. The uninformed traders have a strictly better strategy by choosing $\lambda_q = \begin{cases} \frac{\lambda_I h}{q_1 L}, & for\ q = q_1 \\ 0, & otherwise \end{cases}$.[32]

We then show that $q^* = 1$. Notice that $s_t^q = \frac{2\sigma p_t \lambda_J}{\sum_{i=q}^{\infty} \lambda_i + \lambda_J}$ decreases in $\sum_{i=q}^{\infty} \lambda_i$. Since the best λ_q that maximizes $\sum_{i=q}^{\infty} \lambda_i$ is $\lambda_q = \begin{cases} \frac{\lambda_I h}{L}, & for\ q = 1 \\ 0, & otherwise \end{cases}$. In this case, we reach the one-layer minimum possible spread of

$$s_t^1 = \frac{2\sigma p_t \lambda_J}{\frac{\lambda_I h}{L} + \lambda_J}. \tag{A.2}$$

Finally, recall that $DVol_t \equiv \lambda_I p_t h + \lambda_J p_t L = \lambda_I v_t + \lambda_J p_t L$ is the total dollar volume per unit of time. We therefore derive Corollary 1 directly. Corollary 1 specifies that the dollar bid–ask spread $s_t^{1,L} = \frac{2\sigma \lambda_J L}{DVol_t} p_t^2$, where the superscript L means lot-driven spread. Because we've proved that $q = 1$ is the equilibrium, we drop the superscript "1" and use s_t^L throughout the paper. ■

A.2. *Proof of Lemma 1 and Proposition 2*

The quoted bid–ask spread at $A_t = p_t + \frac{s_t^L}{2} + \left[\Delta - mod\left(p_t + \frac{s_t^L}{2}, \Delta\right)\right]$ and $B_t = p_t - \frac{s_t^L}{2} - \left[\Delta - mod\left(p_t - \frac{s_t^L}{2}, \Delta\right)\right]$ is competitive because any quote that improves these bid and ask prices would lose money. In this proof, we calculate the average widening effect in two steps. First, we show that, under our Poisson jump process, p_t converges to a lognormal

[32] Reducing all child orders to size q_1 reduces transaction costs through two channels. First, uninformed traders no longer walk up the book and pay higher spreads $s_t^2, \ldots, s_t^q > s_t^1$. Second, slicing orders into smaller sizes increases order arrival rates, so the competitive market maker quotes a narrower s_t^1.

distribution, and the residual $mod(p_t, \Delta)$ is asymptotically uniformly distributed within the tick. Second, we solve $s_t^{\Delta} = B_t - A_t - s_t^L$ and show that the uniform distribution leads to an average widening effect of Δ, so the tick-constrained spread is one tick wider than the continuous case in expectation.

First, observe the process through which v jumps up or down by $\sigma\%$ following a Poisson process with intensity λ_J. We then have

$$v_t = v \cdot (1+\sigma)^u \cdot (1-\sigma)^d, \tag{A.3}$$

where $u \sim Poisson\left(\frac{\lambda_J t}{2}\right)$ and $d \sim Poisson\left(\frac{\lambda_J t}{2}\right)$. Taking the log on both sides, we have

$$log(v_t) = log(v) + u \cdot log(1+\sigma) + d \cdot log(1-\sigma). \tag{A.4}$$

When the jump has occurred sufficiently many times, we apply the central limit theorem to (A.4) and $log(v_t)$ converges in distribution to a normal distribution with mean $\mu(t) = log(v) + \frac{\lambda_J t}{2} \cdot log(1+\sigma) + \frac{\lambda_J t}{2} \cdot log(1-\sigma)$ and variance $\Phi(t) = \left(\frac{\lambda_J t}{2} log(1+\sigma)\right)^2 + \left(\frac{\lambda_J t}{2} log(1-\sigma)\right)^2$. Then, v_t follows the lognormal distribution $\mathcal{LN}(\mu(t), \Phi(\mathrm{t}))$, and $p_t = \frac{v_t}{h}$ follows the lognormal distribution $\mathcal{LN}\left(\frac{\mu(t)}{h}, \frac{\Phi(t)}{h^2}\right)$.

Next, we estimate the maximum fluctuation of the probability distribution function within a tick. Let $g(p)$ be the probability distribution function of the lognormal distribution. We compare $g\left(p+\frac{\Delta}{2}\right)$ and $g\left(p-\frac{\Delta}{2}\right)$ and show that, for any $p \gg \Delta$, the relative difference $\left|\frac{g\left(p+\frac{\Delta}{2}\right)-g\left(p-\frac{\Delta}{2}\right)}{g(p)}\right|$ is on the order of $\frac{\Delta}{p}$. With this estimation, the residual of p within a tick is almost uniformly distributed.

Because $p \gg \Delta$, we have $g\left(p+\frac{\Delta}{2}\right) - g\left(p-\frac{\Delta}{2}\right) \approx \Delta g'(p)$, and $\left|\frac{g(p+\frac{\Delta}{2})-g(p-\frac{\Delta}{2})}{g(p)}\right| \approx \left|\frac{\Delta g\prime(p)}{g(p)}\right|$. Inserting the pdf of the lognormal distribution $\mathcal{LN}\left(\frac{\mu(t)}{h}, \frac{\Phi(t)}{h^2}\right)$ into $g(p)$, we have:

$$\left|\frac{\Delta g\prime(p)}{g(p)}\right| = \frac{\Delta}{p}\left(1 + \frac{log(p)-\mu(t)/h}{\Phi(t)/h^2}\right). \tag{A.5}$$

When $t \to \infty$, $\Phi(t)$ goes to infinity on the order of t^2, and $\frac{log(p)-\mu(t)/h}{\Phi(t)/h^2}$ becomes negligible. Thus, for any $p \gg \Delta$, the relative difference $\left|\frac{g\left(p+\frac{\Delta}{2}\right)-g\left(p-\frac{\Delta}{2}\right)}{g(p)}\right|$ is on the order of $\frac{\Delta}{p}$, which is small. The difference is greatest when p is the smallest. As stock exchanges usually delist a stock when its price is consistently lower than \$1.00, $\left|\frac{g\left(p+\frac{\Delta}{2}\right)-g\left(p-\frac{\Delta}{2}\right)}{g(p)}\right| \approx \frac{\Delta}{p} = 10^{-2}$ if $\Delta =$ \$0.01). For a median \$35 stock, the maximum range is even smaller, at $\frac{1}{3500}$, and mostly negligible, and p_t is almost equally likely to lie at \$35.0001 and \$35.0099.[33]

We now solve s_t^{Δ} as a function of $mod(p_t, \Delta)$ and $s_t^L \equiv a\Delta + b$, where $a = 0,1,2,3, ...$ and $b = mod(s_t^L, \Delta)$. We consider breakpoints where $p_t \pm \frac{s_t^L}{2}$ coincides with the tick grids because those breakpoints are boundary cases between "lucky" and "unlucky" scenarios. When a is an even number, $p_t - \frac{s_t^L}{2}$ coincides with a tick grid when $mod(p_t, \Delta) = \frac{b}{2}$, and $p_t + \frac{s_t^L}{2}$ coincides with a tick grid when $mod(p_t, \Delta) = \Delta - \frac{b}{2}$. For any $mod(p_t, \Delta) \in \left[\frac{b}{2}, \Delta - \frac{b}{2}\right]$ (the "lucky" case), the continuous-pricing bid–ask spread is confined within $a + 1$ ticks. Otherwise, the "unlucky" case arises. When a is an odd number, $p_t - \frac{s_t^L}{2}$ coincides with a tick grid when $mod(p_t, \Delta) = \frac{\Delta}{2} + \frac{b}{2}$, and $p_t + \frac{s_t^L}{2}$ coincides with a tick grid when $mod(p_t, \Delta) = \frac{\Delta}{2} - \frac{b}{2}$. For any $mod(p_t, \Delta) \in \left[0, \frac{\Delta}{2} - \frac{b}{2}\right) \cup \left[\frac{\Delta}{2} + \frac{b}{2}, \Delta\right)$ (the "lucky" case), the continuous pricing bid–ask spread is confined within $a + 1$ ticks. Otherwise, the "unlucky" case arises.

The last step is to show that the widening effect is one tick with a uniformly distributed

[33] In principle, any differentiable $g(p)$ with a bounded $g'(p)$ would lead to an approximately uniform distribution within a tick, as long as the variation in p is much wider than a tick so that in any neighborhood of a specific p, $g(p)$ does not exhibit wide variation or a concentrated mass (Anshuman and Kalay 1998). This is arguably the case for all NYSE and NASDAQ listed stocks where the tick size is at most one hundredth of the stock price.

$mod(p_t, \Delta)$. The probability that "lucky" cases arise is its interval length divided by Δ. For both odd and even a, the interval length is $\Delta - b$, so the probability is $\frac{\Delta-b}{\Delta}$. The probability that the "unlucky" scenario arises is then $\frac{b}{\Delta}$. The widened spread is $[(a+1)\Delta - (a\Delta + b)] = \Delta - b$ in "lucky" cases and $2\Delta - b$ in "unlucky" cases. We have $\mathbb{E}(s_t^{\Delta}) = \frac{\Delta-b}{\Delta} \cdot (\Delta - b) + \frac{b}{\Delta} \cdot (2\Delta - b) = \Delta$. ■

A.3. *Proof of Proposition 3*

If we define h^* as shares outstanding under the optimal p^*, Equation (14) becomes

$$p^* = \sqrt{\frac{\lambda_I v \Delta}{2\sigma\lambda_J L}} = \sqrt{\frac{\lambda_I p^* h^* \Delta}{2\sigma\lambda_J L}} \Rightarrow p^* = \frac{\lambda_I h^* \Delta}{2\sigma\lambda_J L}. \tag{A.6}$$

Recall that in Equation (7) $s_t^L = \frac{2\sigma\lambda_J L p_t}{\lambda_I h + \lambda_J L}$. Inserting (A.6) into Equation (7), the expected lot-driven spread under optimal pricing is

$$\mathbb{E}(s_t^L) = \frac{2\sigma\lambda_J L}{\lambda_I h^* + \lambda_J L} \mathbb{E}(p_t) = \frac{2\sigma\lambda_J L}{\lambda_I h^* + \lambda_J L} p^* = \frac{\lambda_I h^* \Delta}{\lambda_I h^* + \lambda_J L}. \tag{A.7}$$

Here $\mathbb{E}(p_t) = p^*$ because p_t is a martingale. Therefore,

$$\mathbb{E}(s^{tot,*}) = \mathbb{E}(s_t^{L,*}) + \mathbb{E}(s_t^{\Delta}) = \frac{\lambda_I \Delta h^*}{\lambda_I h^* + \lambda_J L} + \Delta = \Delta \cdot \left(1 + \frac{\lambda_I h^*}{\lambda_I h^* + \lambda_J L}\right). \tag{A.8}$$

When $h^* \gg L$, we have $\mathbb{E}(s^{tot,*}) \approx 2\Delta$. ■

A.4. *Proof of Corollary 2 and Corollary 3*

Observing a time-weighted average nominal spread $\mathbb{E}(s_t^{tot})$, its lot-driven component $\mathbb{E}(s_t^L) = \mathbb{E}(s_t^{tot}) - \Delta$ will change to $\frac{\mathbb{E}(s_t^{tot}) - \Delta}{H^2}$ following the Square Rule, and the tick-driven component remains Δ. Therefore, we predict that the ex-post nominal spread is

$\frac{\mathbb{E}(s_t^{tot})-\Delta}{H^2} + \Delta$.[34] The nominal price also changes from p_t to p_t/H, so the percentage spread $\frac{\mathbb{E}(s_t^{tot})}{p_t}$ will change to $\frac{(\mathbb{E}(s_t^{tot})-\Delta)/H^2+\Delta}{p_t/H}$. We have

$$\frac{(\mathbb{E}(s_t^{tot})-\Delta)/H^2+\Delta}{p_t/H} = \frac{\mathbb{E}(s_t^{tot})-\Delta}{p_t} \cdot \frac{1}{H} + \frac{\Delta}{p_t} \cdot H \geq 2\sqrt{\frac{\mathbb{E}(s_t^{tot})-\Delta}{p_t} \cdot \frac{\Delta}{p_t}}. \tag{A.9}$$

The equality holds only when

$$\frac{\mathbb{E}(s_t^{tot})-\Delta}{p_t} \cdot \frac{1}{H} = \frac{\Delta}{p_t} \cdot H \Rightarrow H^* = \sqrt{\frac{\mathbb{E}(s_t^{tot})-\Delta}{\Delta}}. \tag{A.10}$$

Therefore, when $h \gg L$, the liquidity-optimal H depends only on the ratio of the observed time-weighted average spread $\mathbb{E}(s_t^{tot})$ and the tick size Δ. ■

A.5. *Proof of Corollary 4*

From Proposition 3, we have $p^* = \sqrt{\frac{\lambda_I v \Delta}{2\sigma\lambda_J L}}$, which is proportional to $\sqrt{\frac{\Delta}{L}}$. Insert p^* into (10), we have

$$\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) = \sqrt{2\sigma v \lambda_I \lambda_J L \Delta}\left(1 + \sqrt{\frac{\lambda_J L \Delta}{8\sigma v \lambda_I}}\right). \tag{A.11}$$

Note that $v = hp$, so the second term is negligible when $h \gg L$. Therefore, when the number of shares outstanding is sufficiently larger than the lot size, the expected transaction cost is proportional to $\sqrt{L\Delta}$. ■

A.6. *Proof of Corollary 5*

In Equation (10), the expected transaction cost $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) = \sigma\lambda_J Lp + \frac{\Delta}{2}\lambda_I \frac{v}{p} +$

[34] Again, we do not need to observe firm fundamentals (σ, v_t, λ_I, and λ_J) to calculate the spread changes caused by a stock split, because the observed spread is a sufficient statistics in determining the split ratio.

$\frac{\Delta}{2}\lambda_J L$ depends on the firm's choice of p. Inserting the proportional lot size $L = \mathbb{L}(p) = k^L/p$, we have:

$$\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) = \sigma\lambda_J k^L + \frac{\Delta}{2}\lambda_I \frac{v}{p} + \frac{\Delta}{2}\lambda_J \frac{k^L}{p}. \tag{A.12}$$

(A.12) indicates that the seemingly flexible proportional lot size imposed a unified dollar lot size k^L on all stocks, and the lot-driven component depends on k^L but not on p. In other words, firms cannot adjust their nominal prices to reduce market makers' adverse-selection costs, and their nominal price choices affect only the relative tick size. Therefore, the expected transaction cost decreases monotonically with p. The proportional lot size essentially removes one side of the tick/lot trade-off and encourages $p \to \infty$.

On the other hand, if we insert the proportional tick size $\Delta(p) = k^\Delta p$ into (10), we have:

$$\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) = \sigma\lambda_J Lp + \frac{k^\Delta}{2}\lambda_I v + \frac{k^\Delta p}{2}\lambda_J L. \tag{A.13}$$

(A.13) indicates that the proportional tick-size system imposed a unified relative tick size k^Δ on all stocks. No firms can reduce their transaction costs below $\frac{k^\Delta}{2}\lambda_I v$. With a uniform lot size and a proportional tick size, the transaction cost increases monotonically with p. The proportional tick size essentially removes the other side of the tick/lot trade-off and encourages $p \to 0$, where $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) = \frac{k^\Delta}{2}\lambda_I v$.

Similarly, when both proportional tick- and lot-size systems are implemented, we have:

$$\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot DVol_t\right) = \sigma\lambda_J k^L + \frac{k^\Delta}{2}\lambda_I v + \frac{k^\Delta}{2}\lambda_J k^L. \tag{A.14}$$

Under the fixed Δ and L system, firms adjust their nominal prices to choose their optimal dollar lot sizes and relative tick sizes. (A.14) shows that the proportional tick and lot system is a one-size-fits-all system: it imposes a homogenous dollar lot size and relative tick size on all stocks. Next, we show that such a system harms liquidity provision if the regulator chooses k^L and k^Δ using any representative stock.

We denote $\chi(p) = \frac{p}{p_\Omega}$ as the distance between the representative price p_Ω and a stock priced at p. For a stock optimally priced at p, its new tick size is χ times Δ, while its new lot size becomes χ^{-1} times L. Insofar as the tick- (lot-) driven percentage spread is proportional to the relative tick (dollar lot) size, the new nominal spread is $(\chi^{-1} + \chi)\Delta$. Observe that $(\chi^{-1} + \chi)\Delta \geq 2\Delta$, where the equality holds only if $p_t = p_\Omega$ because its relative tick size and dollar lot size are unchanged. The bid–ask spread widens for all stocks, with $p_t \neq p_\Omega$.■