THE OPTIMAL NOMINAL PRICE OF A STOCK: A TALE OF TWO DISCRETENESSES

By SIDA LI AND MAO YE *

Economists commonly assume that price and quantity are continuous variables, while in reality both are discrete variables. As U.S. regulation mandates a one-cent minimum tick size and a 100-share minimum lot size, we predict that less volatile stocks and more active stocks should choose higher prices to make pricing more continuous and quantity more discrete. Despite heterogeneous optimal prices, all firms achieve their optimal prices when their bid–ask spreads equal two ticks, 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 stock liquidity. We find that most stock splits move the bid–ask spread closer to two ticks and that correct splits contribute 94 bps to split announcement returns. Optimal pricing can increase median U.S. stock value by 106 bps and total U.S. market capitalization by $93.7 billion.

* Li: University of Illinois at Urbana-Champaign (email: sidali3@illinois.edu); Ye: University of Illinois at Urbana-Champaign and NBER, 340 Wohlers Hall, 1206 S. 6th Street, Champaign, IL, 61820 (e-mail: maoye@illinois.edu). We thank James Angel, Kerry Back, Malcolm Baker, Dan Bernhardt, Hendrik Bessembinder, Kevin Crotty, John Campbell, Charles Cao, Xavier Gabaix, Robin Greenwood, Albert (Pete) Kyle, Giang Nguyen, Maureen O'Hara, Gideon Saar, Andrei Shleifer, Jeremy Stein, Bart Zhou Yueshen, 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, 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.

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 is 100 shares for most stocks.[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.

Panel A in Figure 1 shows that stocks are most liquid when their prices are neither too high nor too low, and Panel B shows preliminary evidence that the trade-off between discrete quantity and discrete pricing drives this U-shaped pattern. Stocks whose prices are too low suffer from tick-size constraints. As the bid–ask spread cannot drop below one cent, the percentage bid–ask spread decreases with prices for stocks whose bid–ask spread is one cent, as indicated by the bottom left frontier in the figure. Stocks whose prices are too high suffer from lot-size constraints. Even for large firms such as Google and Amazon, the median depth at the National Best Bid and Offer (NBBO) and trade sizes are exactly 100 shares. Once the lot size is binding, an increase in share prices amplifies the market maker's obligation to maintain round lots; the percentage bid–ask spread tends to increase with prices because market makers lose more money once they are adversely selected. Also, Figure 1 indicates that the optimal price depends on characteristics: the optimal price seems to increase with a stock's market cap. What, then, is the optimal price that maximizes liquidity? Which factors determine the optimal price? How large are the economic benefits gained from managing the nominal price? Our paper provides both theoretical and

[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.

empirical answers to these questions.

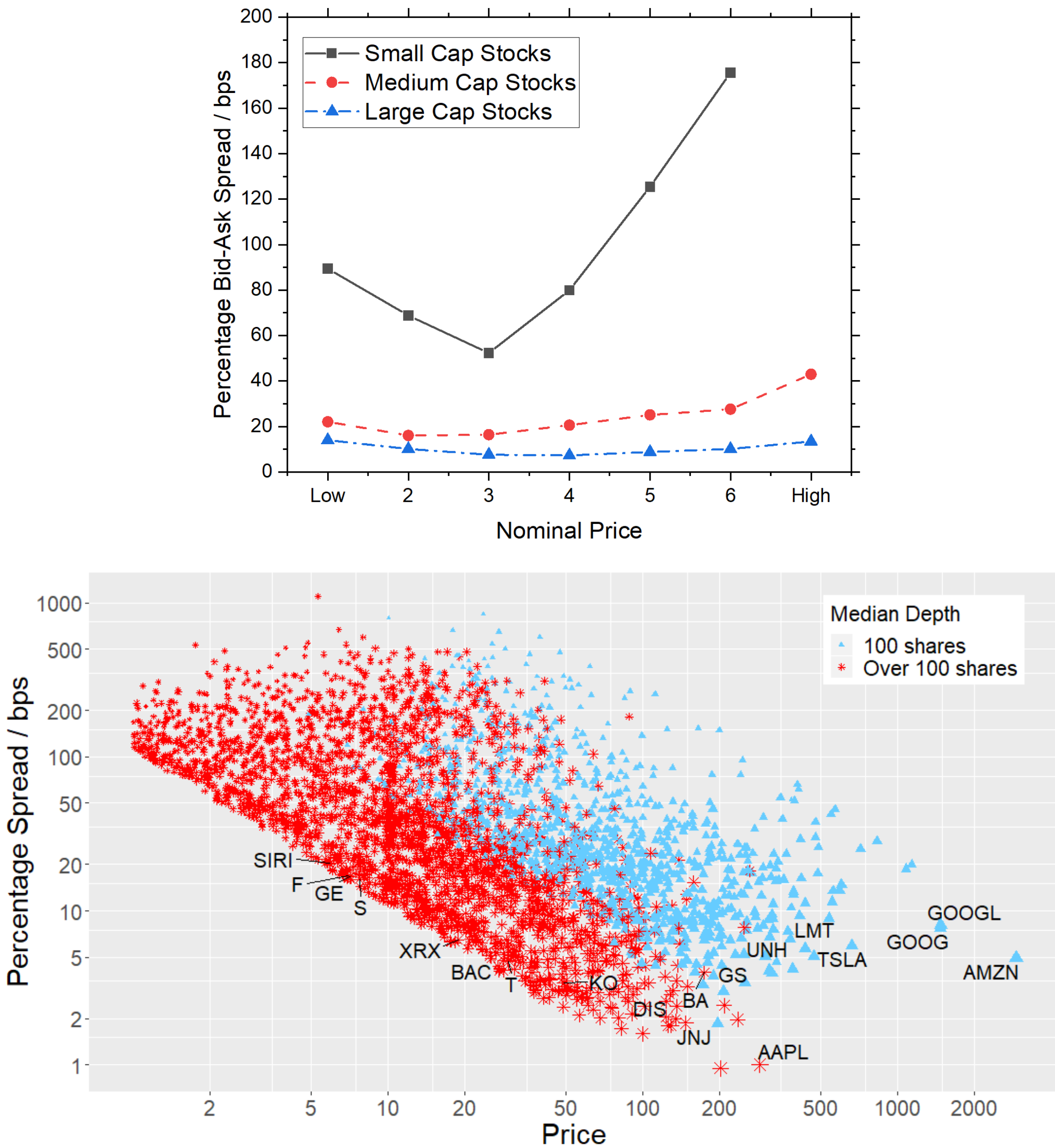

FIGURE 1.—U-Shaped relationship between liquidity and prices: These figures illustrate the relationship between percentage spreads and nominal prices. Our sample includes all U.S.-listed common stocks that that are subject to the 1-cent tick size, a 100-share lot size, and at least a $1 nominal price. For Panel A, we take the average spread across price baskets and group stocks by market caps. The lines consisting of squares, circles, and triangles represent small-, medium-, and large-cap stocks, respectively. Price baskets are selected such that each basket contains a similar number of stocks. In Panel B, we plot each firm as a triangle or asterisk, where larger shapes represent larger market-cap firms. Blue triangles represent stocks with 100 shares at the median national best bid/offer (NBBO) depth while red asterisks represent stocks with median NBBOs of more than 100 shares. The bottom-left boundary represents the 1-cent tick size constraint.

We discover a Two-Tick Rule for optimal pricing. Theoretically, we predict that firms should seek heterogeneous pricing based on their volatility and dollar volumes, but all firms reach their optimal prices when their bid–ask spreads are two ticks wide. Intuitively, as the friction caused by a discrete price equals one tick, all firms reach their optimal prices when the friction caused by the discrete lot also equals one tick, i.e., when their bid–ask spreads equal two ticks. Our empirical results support the two-tick prediction. Stock splits improve liquidity if they move the nominal spread towards two ticks, whereas those that move the spread away from two ticks reduce liquidity. We find that most stock splits move the nominal spread towards two ticks. As a result, stock splits reduce average percentage drops in the bid–ask spread by 15.22 bps. The liquidity gains obtained from stock splits generate a 94-bps increase in firm value. We estimate that the median U.S. stock value would increase by 106 bps if all firms were to move to their optimal prices.

Our paper starts with a three-stage model. In the first stage, a regulator chooses tick and lot sizes. In the second stage, a firm sets its share price such that it minimizes the expected transaction cost for its traders. In the third stage, the market opens, where the stock is traded by three types of agents. A market maker posts competitive bid prices to sell and ask prices to buy, and her quote size cannot be smaller than one lot. Uninformed traders, who arrive with exogenous needs to buy or sell a security, choose how to divide their demand. For example, an uninformed trader can submit all his demand in one order or break his demand into a series of child orders of the minimum lot size. Informed traders know the value of a security before each jump, and they profit from adversely selecting the market maker. The market maker earns the bid–ask spread if an uninformed trader hits her quotes, but she loses money when an informed trader adversely selects her quotes.

In Section 3, the regulator in our model mandates a discrete quantity but keeps pricing continuous. The main result reported in Section 3 is the Square Rule: an H-fold reduction in share price leads to an H^2-fold reduction in the bid–ask spread and thereby an H-fold drop in the percentage spread. A lower price increases liquidity because of traders' interactions in the third stage. To minimize the loss in price caused by informed traders, the market maker always displays a minimum lot and refills the lot once it is consumed.

Uninformed traders also break their demand into series of child orders of the minimum lot size. As the lot size is binding, an H-fold reduction in share price leads to an H-fold reduction in the loss for its market maker. In turn, the market maker can afford a percentage spread that is H times tighter or a bid–ask spread that is H^2 times tighter. A discrete quantity, therefore, favors a lower price per share.

We report our main theoretical results in Section 4, where both price and quantity are discrete. When pricing becomes discrete, the competitive ask (bid) prices are the lowest (highest) prices above (below) the break-even ask (bid) prices. We find that the expectation of the widening effect is one tick for any break-even spread. This result decomposes bid–ask spread s into two components: a tick-driven component (Δ) and a lot-driven component ($s - \Delta$). The tick-driven component Δ favors a higher price and the lot-driven component favors a lower price. The trade-off between these components leads to the two-tick rule: every stock reaches its optimal price when its bid–ask spread equals two ticks. The same two-tick bid–ask spread leads to heterogeneous optimal pricing. A volatile stock should choose low prices because greater volatility increases adverse selection risk and the firm should reduce its share price to reduce the loss its market makers experience. Holding volatility fixed, a stock whose dollar volume is higher, either because of its larger market cap or its higher turnover rate, should choose higher prices because the stock enjoy lower percentage spreads, making discrete pricing the main friction.

Although a firm's optimal prices depend on its fundamental characteristics, the Modified Square Rule in Section 4 indicates that a firm does not need to calibrate these characteristics to achieve its optimal price beyond observing its current bid–ask spread. When a stock splits, its tick-driven spread remains the same, but its lot-driven spread follows the square rule. The Modified Square Rule predicts that an H-for-1 split leads to a bid–ask spread of $\Delta + \frac{s-\Delta}{H^2}$. In turn, a firm can choose H such that $\Delta + \frac{s-\Delta}{H^2} = 2\Delta$ to achieve its optimal price. When $\Delta = 1$ cent, the optimal H is $\sqrt{s-1}$.

In Section 5, we allow the regulator to change tick and lot sizes. We derive the Square Root Rule. Suppose that the regulator permanently increases the tick size from one cent to five cents. The firm's optimal response is a $\sqrt{5}$-for-1 reverse-split, which maintains the same contributions of the tick size ($\sqrt{5}$) and the lot size ($\sqrt{5}$). The Square-Root Rule leads to a spillover effect: a policy initiative that makes pricing more discrete would, in equilibrium, make quantity more discrete. The Two-Tick Rule still holds, but two ticks are now ten cents, leading to a $\sqrt{5}$ increase in the percentage spread. The Square Root Rule also applies to lot size. The best response to a 100-fold reduction in lot size is 10-for-1 reserve splits, which creates a tenfold drop in the dollar lot size, the relative tick size and the percentage spread. Thus, we encourage the SEC to consider reducing tick and lot sizes to improve market liquidity.

In Section 5, we also allow the regulator to switch from uniform tick and lot sizes to proportional tick and lot sizes. We find that proportional tick and lot sizes 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 an optimal price to balance discrete prices and quantities. The proportional system destroys this degree of freedom and reduces liquidity if the regulator uses any existing stock as the benchmark against which to jumpstart the proportional system. For example, consider a \$300 stock and a \$3 stock. Both have chosen the optimal two-cent spread, one cent from the tick and one cent from the lot. A proportional system based on a \$30 stock will assign a tenfold wider tick size and a tenfold smaller lot size to the \$300 stock, leading to a 10.1 cent spread ($= 10 + 0.1$). On the other hand, the \$3 stock also loses its optimal tick–lot balance and achieves a bid–ask spread of 10.1 cents ($= 0.1 + 10$). Therefore, the proportional system harms liquidity because it imposes a uniform level of discreteness on stocks with heterogeneous characteristics.

In Section 6, we test our empirical predictions in the cross-section. For a given nominal price, the Modified Square Rule predicts a firm's liquidity, no matter whether the firm's price is optimal. In the cross-section, the Modified Square Rule explains 81% of the variation in the bid–ask spread with only three variables modeled by our paper (price, dollar volume, and volatility). Our three-factor model of liquidity outperforms existing benchmarks (Madhavan 2000; Stoll 2000) even though we use only a subset of their explanatory variables. The main driver of this outperformance is the functional form. Madhavan (2000) controls for $price^{-1}$ while Stoll (2000) controls for $log(price)$. 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 our specification: subtracting one tick from the bid–ask spread to control for the tick-driven spread and then taking the log of the price to control for the lot-driven spread.

Our model also helps to remove redundant variables that have been included in previous specifications. For example, all existing specifications control for the market cap, following the intuition that large stocks are more liquid. Our model suggests that the market cap affects liquidity through its impact on dollar volumes. Holding the share-turnover rate fixed, a large-cap stock is more liquid. Our model predicts that a small firm with higher turnover should, however, be as liquid as a large firm as long as their dollar volumes are the same. Our interpretation addresses a puzzle raised in Stoll (2000), who finds that the regression coefficient before the market cap is not always positive after controlling for the dollar volume. The 81% R^2 in our regression suggests that future empirical research may use our three-factor model of liquidity as a benchmark to evaluate additional liquidity predictors.

Our theoretical model also implies a two-factor model of share prices, and we find that volatility and dollar volume explain 57% of the cross-sectional variation in share prices. Our paper rationalizes several puzzles that have been documented in the behavioral finance literature. Baker, Greenwood, and Wurgler (2009) find it puzzling that volatile firms are more likely to split their stocks because these firms have a "greater chance of reaching a low price anyway." Shue and Townsend (2018) provide a behavior-based interpretation. As investors think in part about stock-price changes in dollar rather than percentage units, low-priced stocks should experience more extreme return responses to news. We provide two alternative explanations for the negative volatility–price relationship. The first interpretation comes from tick constraints. Stocks whose volatility is higher face higher adverse selection risk and thereby higher percentage bid–ask spreads. Higher percentage bid–ask spreads then relieve tick constraints and provide stronger incentives for firms to choose lower prices. The second interpretation comes from lot constraints. A rise in volatility increases market makers' losses when they are adversely selected, and firms should choose lower prices to reduce the adverse-selection risk. Weld et al. (2009) find that firms choose prices that are similar to those chosen by similarly sized firms and industrial peers, and they conjecture that social norms drive such clustering. We find that the explanatory power of industry fixed effects is superseded by that of volatility. Therefore, firms in the same industry may choose similar prices because their volatilities are similar. Our model also rationalizes the cluster of nominal prices around size, as an increase in size or turnover increases dollar volume. An increase in dollar volume reduces the percentage spread and makes tick size a more binding constraint. Therefore, our paper not only rationalizes price clustering but also explains why large stocks cluster at higher prices.

In Section 7, we test the empirical predictions using stock splits. First, we find that changes in bid–ask spreads after stock splits match almost exactly with what our theoretical model implies: a 1 bps increase in the percentage spread predicted by the Modified Square Rule leads to a 1.02 bps increase in the realized change. Second, we find that 1,077 of 1,196 splits move the bid–ask spread closer to the two-tick optimum. Therefore, most splits are correct. Among the 107 incorrect splits, 74 make the correct decision to split, except that they choose a split ratio that is overly aggressive. Overall, the percentage spread drops by 15.22 bps after splits. The effect on liquidity is so significant that it affects firm value. We find that a 1 bps reduction in the predicted percentage spread increases firm value by 6.18 bps. Therefore, correct split ratios increase firm value by 94 bps, which is more than one-third of the split-announcement return of 273 bps.

Although we show that the nominal prices and stock splits that firms choose are overall rational, firms may still end up with suboptimal prices. Section 8 provides back-of-the-envelope calculations for the benefit of moving towards optimal prices. The most salient evidence comes from firms that are similar in fundamentals but choose drastically different prices. For example, Amazon's stock is 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 is 1.95 cents (0.77 bps). We find that nominal prices explain this 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, which is similar to Microsoft's spread. Amazon holders paid $684 million in the (half) bid–ask spread per year, but Apple holders paid only $60 million. The split ratio to achieve the optimal two-tick spread for Amazon is $\sqrt{152}$-for-1, which would save Amazon shareholders $574 million per year in transaction costs. Amazon's market cap would increase by $3.96 billion. Overall, the market value of U.S. firms would increase by $93.7 billion if all firms were to move to their optimal prices.

As the first study to examine discreteness in both pricing and quantity, our paper offers significantly different predictions than are implied by models where only one variable is continuous. Angel's (1997) optimal tick-size hypothesis focuses only on price discreteness, and he predicts that 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. Although a 2-for-1 reverse split retains the original relative tick size, it also doubles the lot-driven spread, leaving the percentage spread unchanged, as if nothing had been done. Budish, Cramton, and Shim (2015, "BCS" hereafter), who consider a market with discrete quantities but continuous pricing, find that public information leads to a positive percentage spread. We find that this percentage spread is a linear function of lot size. Therefore, the percentage spread converges on zero when the lot size converges on zero. Holding lot size fixed, a firm can also reduce the dollar lot size and its percentage spread through aggressive stock splits when pricing is continuous. In our model, the economic factor that prevents such aggressive splits is discrete pricing.

We address two questions in the corporate finance literature. 1) Why do firms split their stocks? 2) What explains the positive returns that follow splits? Our tick-and-lot channel helps us rationalize the puzzles raised by two canonical channels for stock splits: the signaling channel and the clientele channel. In the signaling channel, the cost of the signal comes from reduced liquidity (Brennan and Copeland, 1988), yet we find that splits increase liquidity. Also, Fama et al. (1969), Lakonishok and Lev (1987), and Asquith, Healy, and Palepu (1989) find that earnings, profits, and stock prices increase significantly before splits but not after splits. Their results do not support the signaling channel but support our tick-and-lot channel. A previous increase in a stock price increases the lot constraint on that stock, and stock splits provide the best response to lot constraints. In the clientele channel (Lamoureux and Poon, 1987 and Maloney and Mulherin, 1992), firms use stock splits to attract retail traders and expect that an increase in the number of uninformed traders increases volume and liquidity, yet we find that institutional holdings increase after splits in our sample.

Our empirical studies focus on cross-sectional variation and event studies around a short time window, because the parameters in our model may change over a long period of time. Nevertheless, the economic drivers underlying our model provide a rationale for some long-term time trends following the proliferation of electronic trading. For example, we provide the first unified explanation of four salient facts that emerged after trading became automated: a reduction in the bid–ask spread (Hendershott, Jones, and Menkveld, 2011), the decline in depth towards one lot (Angel, Harris, and Spatt, 2015), the dominance of one-lot trades (O'Hara, Yao, and Ye, 2014), and the proliferation of algorithmic traders who are not as fast as high-frequency traders (HFTs) (O'Hara, 2015). We generate these predictions in one model because we model interactions between distinct types of algorithmic traders, whereas most studies include at most one type of algorithmic trader: HFTs.[2] Note that liquidity demanders' execution algorithms do not need to execute as quickly as those of HFTs. They just need to be fast enough to slice and dice large latent demand into one-lot pieces. Also, one of the most vexing puzzles in the literature on nominal prices and tick size involves understanding why firms do not split 1-to-6.25 after the 2001 decimalization standard reduced 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 the decimalization. Most importantly, our model predicts that the optimal price will be the same if tick size and lot size reduced by

[2] A notable exception is Li, Wang, and Ye (2021), who study competition in liquidity provision between HFTs and slower execution algorithms.

the same amount. The proliferation of electronic trading allows traders to slices 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 years, we see an increase in nominal prices and a lack of stock splits over the past two decades. We summarize these time trends in Figure 2.

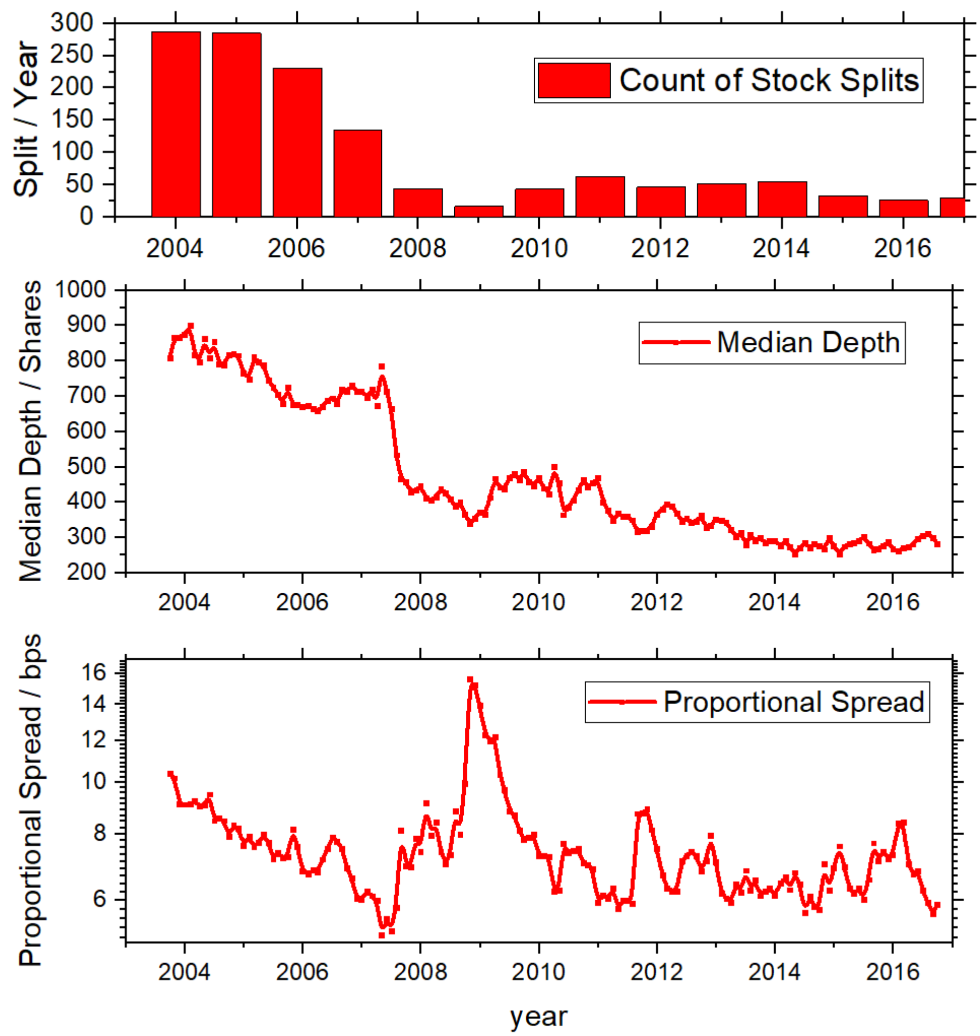

FIGURE 2.—Simultaneous reduction of the proportional spread, market depth, and the number of stock splits: This figure shows the time trend in stock splits, the median depth, and the percentage spread from 2003 through 2016. Each dot represents a monthly observation. Our sample includes all NYSE-listed common stocks that have a 1 cent tick size, a 100-share lot size, and at least a $1 nominal price throughout a given month. We calculate the monthly mean of the depth and spread for each stock and take the median across stocks.

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 -2. We do not model the regulator's utility function because our main purpose is to examine the firm's and traders' best responses under varying tick- and lot-size regimes. 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 incorporate regulations that apply in other countries.

***Stage 2: The firm's decision*:** The firm in our model starts at value v at time -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 λ_J. The size of the jump is σv_t or $-\sigma v_t$ at equal probability, so v_t is 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 $t \in (0, +\infty)$.

***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.

---

[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 (odd lots) in size.

Uninformed traders arrive at a Poisson process at intensity λ_I. Each uninformed trader has inelastic demand to buy or sell a fraction of the firm at equal probability. For simplicity in notation, we normalize the fraction to 1 such that each uniformed trader's liquidity demand is h shares. We call the demand of h shares parent orders. Uninformed traders 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 variable $f(q)$, where $f(q)$ is the frequency of child orders of q lots. The total size of all child orders equals the size of the parent order, that is, $h = \sum_{q=1}^{h/L} qLf(q)$. For example, uninformed traders can slice all child orders to the minimum lot, such that all orders are in minimum lot size L. Then,

$$f(q) = \begin{cases} \frac{h}{L}, & for\ q = 1 \\ 0, & otherwise \end{cases}.$$

In this case, the arrival rate for orders of the minimum lot is $\lambda_I \frac{h}{L}$, whereas the arrival rate for orders of all other sizes is 0. The other extreme strategy is to submit the total demand as one large order, that is,

$$f(q) = \begin{cases} 1, & for\ q = \frac{h}{L} \\ 0, & otherwise \end{cases}$$

In this case, the arrival rate for the order of size h is $\lambda_I \cdot 1$, whereas the arrival rate of all other sizes is 0. In general, when uninformed traders choose $f(q)$, they effectively choose an order arrival rate of size q as $\lambda_I f(q)$. The uninformed traders choose $f(q)$ to minimize the total transaction costs of their parent orders.

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 bids above the fundamental value or all asks below the fundamental value. 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.[4] All other firm fundamentals, σ, λ_I, and λ_J, are public information for traders and the firm.

After observing $f(\cdot)$, the market maker quotes competitive price on the bid and ask side of the market. Her choice variable is a set of bid and ask prices $\{A_t^q, B_t^q\}$, where q stands for the price for the q^{th} share. A_t^q and B_t^q 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 chooses a tick size of $\Delta = 0$ and a discrete lot size of $L > 0$. We solve the model through backward induction. In subsection 3.1 we solve traders' optimal choices 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 in Stage 2.

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 exactly one lot each. The market maker always displays one lot at the 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 to child orders with heterogeneous sizes. If not,

[4] 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.

suppose uninformed traders choose two sizes, q_1 and q_2 ($q_1 < q_2$).[5] 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 with 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 with order size q_2. Therefore, q_2 child orders execute at worse prices than q_1 child orders because the former walks 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 homogenous child order size, only child order sizes of one minimum lot 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 choose to slice parent orders into child orders of two lots. Then, the competitive market maker needs to maintain a quote for two lots. 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 the result is that a decrease of child order size q_1 increases $f(q_1)$. Therefore, the arrival rate of uninformed child orders reach its maximum ($\lambda_I \frac{h}{L}$) 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 breakeven 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.[6] Brogaard et al. (2015) document the existence of "SlowColos" who co-locate at a stock exchange but are slower

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

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 reduce adverse-selection risk or to select other traders adversely.

The next step in solving the equilibrium is to pin down the transaction cost, measured by 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 $\frac{\mathcal{S}_t^L}{2}$ per dollar. 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 he is sniped. Therefore, the equilibrium percentage spread solves

$$\lambda_I p_t h \cdot \frac{\mathcal{S}_t^L}{2} = \lambda_J p_t L \cdot (\sigma - \frac{\mathcal{S}_t}{2}). \tag{1}$$

The solution is

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

The equivalent nominal bid–ask spread is

$$s_t^L \equiv \mathcal{S}_t^L p_t = \frac{2\sigma\lambda_J L}{\lambda_I h + \lambda_J L} p_t. \tag{3}$$

Proposition 1. (Continuous pricing bid–ask spread) *With zero tick size and lot size L, 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 L shares of liquidity at $p_t \pm \frac{s_t^L}{2}$ *and she refills L shares around* p_t *after each trade.*

(ii) Uninformed traders slice their demand into a series of child orders. Each child order includes L shares.

(iii) Informed traders adversely select L shares 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. The policy and market solutions are economically equivalent under continuous pricing. Effectively, both solutions reduce the market maker's adverse-selection risk by making quantity more continuous. In Section 4 and 5, we show that policy and market solutions are no longer equivalent when price 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 the equation by p_t, we have

$$s_t^L = \frac{2\sigma\lambda_J L p_t^2}{\lambda_I p_t h + \lambda_J p_t L}. \tag{4}$$

Notice that the denominator is equal to the dollar volume per unit of time. Denote

$$DVol_t \equiv \lambda_I p_t h + \lambda_J p_t L = \lambda_I v_t + \lambda_J p_t L. \tag{5}$$

Then we 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 the 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 price, leaving the percentage spread unchanged. Another p_t-time increase comes from the increase in adverse-selection risk, as the competitive market maker incurs higher costs for sustaining one-round-lot liquidity at the BBO. 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. 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

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

Notice that minimizing the expected transaction cost is economically equivalent to minimizing the expected percentage spread. In fact, when the price is continuous, they 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 $\frac{\Delta}{p_t}$ and $\mathbb{E}(\frac{\Delta}{p_t})$ does not have an analytical form.

The firm's objective function under continuous pricing is very intuitive: $\sigma\lambda_J L \cdot \frac{v}{h}$ is the market maker's expected adverse-selection cost per unit of time. Thus, the firm's objective function is to minimize its market maker's adverse-selection cost. A decrease of 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 with more trades at the smaller dollar lot size. Under continuous pricing, firms should choose $h \to \infty$ and $p \to 0$. The result is intuitive. When the lot size is the only friction, firms should split their stocks aggressively to minimize such friction. 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 results 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 caused 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 solve traders' optimal decisions given h, L, and Δ, and we quantify the frictions generated by the discrete tick size. 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 if the firm is traded at suboptimal price.

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. This is because more aggressive quotes lose money while less aggressive quotes are not competitive. Therefore, discrete pricing widens the bid–ask spread.[7]

[7] Notice that s_t^L is the break-even spread for the first lot of liquidity. As discrete price widens the spread, the competitive market maker may quote more than one lot of liquidity at A_t and B_t. 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. The analytical solution of equilibrium depth quoted by the market maker can be solved through a Markov transition model in the spirit of Li, Wang, and Ye (2021) and is available upon request. We do not characterize the depth in this model because depth does not affect outcomes in equilibrium. The uninform traders never walk the book in equilibrium and they always buy at A_t and sell at B_t. Therefore, A_t and B_t provide sufficient statistics to calculate the uninformed traders' transaction costs.

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 position of p_t within the tick grids, i.e., $mod(p_t, \Delta)$. More importantly, Proposition 2 shows that the expectation of the widening effect is one tick.

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* $\mathbb{E}(s_t^{\Delta}) = \Delta$.

The intuition for the one-tick average widening effect is as follows. For a Poisson jump process, the distribution of 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. Therefore $mod(p_t, \Delta)$ converges to uniform distribution because any residual value within the tick is equally likely.[8]

[8] This argument is generally true for any distribution with a smooth, non-clustering probability density function. 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

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 3 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, 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 left, 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 ($s_t^\Delta = 2\Delta - b$) in magnitude.

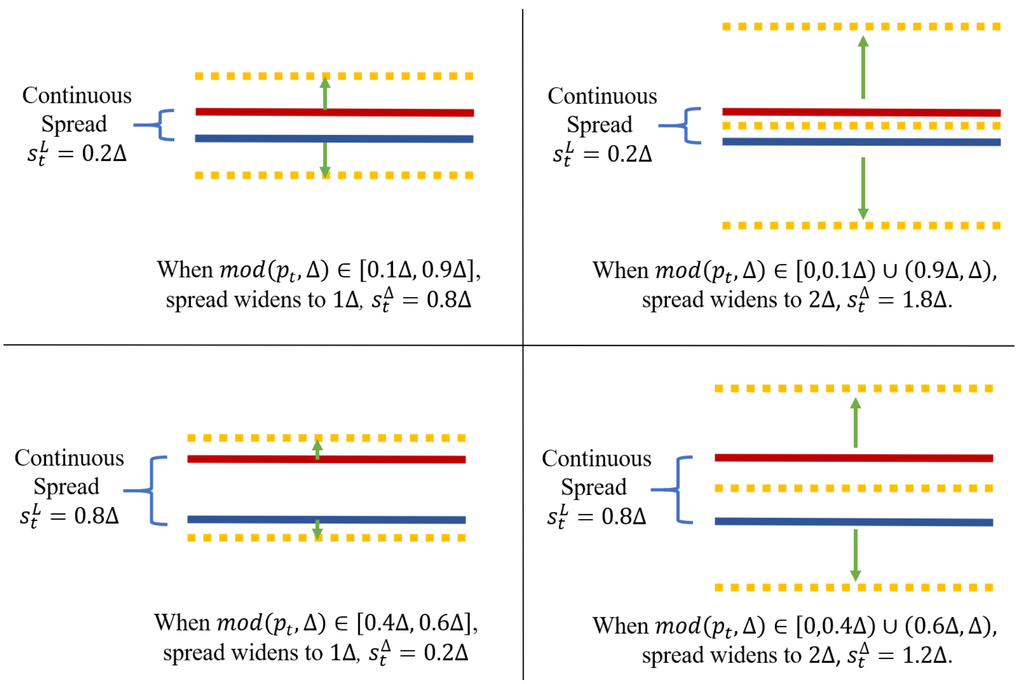

FIGURE 3. —Average widening effect of one tick: This figure illustrates the bid–ask widening effect. The red solid upper bars are the ask prices under continuous pricing and the blue solid lower bars are the bid prices under continuous pricing. The yellow dots represent the tick grids. The green arrows 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 1 tick, and the two panels on the right illustrate the "unlucky" cases where the bid–ask spread widens by more than 1 tick.

---

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.

Proposition 2 shows that the expected widening effect for any break-even spread s_t^L is one tick. Figure 3 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 in 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 $b = 0.8\Delta$, $p_t \pm \frac{s_t^L}{2}$ 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

$$\mathbb{E}\left(s_t^{\Delta}\right) = \frac{\Delta - b}{\Delta} \cdot (\Delta - b) + \frac{b}{\Delta} \cdot (2\Delta - b) = \Delta. \tag{6}$$

Equation (6) implies that we can decompose the bid–ask spread under discrete pricing, s_t^{tot}, into a lot-driven component s_t^L and a tick-driven component s_t^{Δ}. The lot-driven component s_t^L is equal to the continuous pricing spread in Section 3, which follows the Square Rule. The tick-driven component equals to Δ in expectation. An increase in the nominal price inflates the lot-driven spread but it dilutes the widening effect (or tick-driven spread) proportionately.

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

$$
\begin{aligned}
\min_p \mathbb{E}\left(\frac{s_t^{tot}}{2p_t}\cdot Dvol\right) &= \min_p \mathbb{E}\left(\frac{s_t^{L}+s_t^{\Delta}}{2p_t}\cdot Dvol\right)\\
&= \min_p \mathbb{E}\left[\left(\frac{\frac{2\sigma\lambda_J L p_t}{\lambda_I h+\lambda_J L}+s_t^{\Delta}}{2p_t}\right)\cdot\left(\lambda_I p_t h+\lambda_J p_t L\right)\right]\\
&= \min_p \mathbb{E}\left[\sigma\lambda_J p_t L+\frac{s_t^{\Delta}}{2}\left(\lambda_I h+\lambda_J L\right)\right]\\
&= \min_p \left[\sigma\lambda_J L\cdot\mathbb{E}[p_t]+\frac{\mathbb{E}(s_t^{\Delta})}{2}\left(\lambda_I h+\lambda_J L\right)\right] = \min_p\left[\sigma\lambda_J Lp+\frac{\Delta}{2}\lambda_I\frac{v}{p}+\frac{\Delta}{2}\lambda_J L\right]
\end{aligned}
\tag{7}
$$

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

$$\sigma\lambda_J Lp+\frac{\Delta}{2}\lambda_I\frac{v}{p}\geq 2\sqrt{\sigma\lambda_J Lp\cdot\frac{\Delta}{2}\lambda_I\frac{v}{p}}. \tag{8}$$

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

$$p^{*}=\sqrt{\frac{\lambda_I\Delta v}{2\sigma\lambda_J L}}. \tag{9}$$

Given the tick and lot sizes, formula (9) shows that the optimal nominal price decreases with volatility and increases with dollar volume. An increase in volatility, caused either by an increase in jump size σ or an increase in jump frequency λ_J, reduces the optimal nominal price. Holding all else equal, an increase in volatility increases adverse-selection risk. Therefore, a firm should choose a lower price to reduce adverse-selection risk for its market makers, until the contribution of the lot size is again equal to the contribution of the tick size. Tick size constraints also provide an intuition for this result. An increase in volatility increases the percentage spread, and the bid–ask spread in dollars becomes less constrained by tick size. Therefore, the firm should choose lower price because of reduced tick size constraints.

Formula (9) also shows that an increase in dollar volume, caused either by an increase in market cap v or an increase in turnover rate λ_I, increases the optimal price. Holding volatility fixed, an increase in market cap or turnover increases investors' liquidity needs, reduces the percentage spread, and makes the discrete price a larger friction for the bid–ask spread. Therefore, firms with larger market caps or higher turnover should choose higher prices.

Proposition 3 shows that heterogeneous optimal prices lead to the same two-tick bid–ask spread. Intuitively, firms choose heterogeneous prices because of their varying fundamental characteristics, but all firms reach optimal prices when the contribution of the tick size equals the contribution of the lot size. Mathematically, this implies a two-tick bid–ask spread under a trivial assumption that a firm's total shares outstanding h is much larger than its lot size L.[9]

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

Proposition 3 indicates that stocks whose average bid–ask spreads are narrower than two ticks are more tightly constrained by the tick size, and those firms can infer that their nominal prices are too low. Stocks whose average bid–ask spreads are wider than two ticks are more strictly lot-bound and their prices are too high. Fortunately, a firm does not need to calibrate σ, λ_J, and λ_I to estimate its optimal price, because its current average bid–ask spread provides sufficient statistics for this decision. We introduce the optimal split ratio

[9] The minimal shares outstanding in our sample is 56,600 round lots.

in the next subsection.

4.3. Optimal Split Ratio: The Modified Square Rule

In this subsection, we introduce the Modified Square Rule, which shows that the average bid–ask spread is a sufficient statistic for the firms to choose the optimal split ratio. Lemma 1 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

$$\mathbb{E}(s_t^{tot}) = \mathbb{E}\left(s_t^L + s_t^\Delta\right) = \mathbb{E}(s_t^L) + \Delta. \tag{10}$$

The lot-driven component follows the Square Rule (Proportion 2). As $s_t^L = \frac{2\sigma\lambda_J L}{\lambda_I h+\lambda_J L} p_t$, and the total shares outstanding $h \gg L$, we have $s_t^L \approx \frac{2\sigma\lambda_J L}{\lambda_I h} p_t$. Similarly, the lot-driven spread after an H-for-1 split is $\frac{2\sigma\lambda_J L}{\lambda_I Hh+\lambda_J L} \cdot \frac{p_t}{H}$. When $h \gg L$, we have

$$s_t^{L,post} \approx \frac{2\sigma\lambda_J L}{\lambda_I h H^2} p_t \approx \frac{s_t^L}{H^2}. \tag{11}$$

Therefore, the expected post-split bid–ask spread is

$$\mathbb{E}(s_t^{tot,post}) = \mathbb{E}(s_t^{L,post} + s_t^{\Delta,post}) \approx \mathbb{E}\left(\frac{s_t^L}{H^2} + s_t^{\Delta,post}\right) = \frac{\mathbb{E}(s_t^L)}{H^2} + \Delta = \frac{\mathbb{E}(s_t^{tot})-\Delta}{H^2} + \Delta. \tag{12}$$

Therefore, 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 $H^* = \sqrt{\frac{\mathbb{E}(s_t^{tot})-\Delta}{\Delta}}$. In Corollary 2, we modify the Square Rule specified in Corollary 1 to accommodate discrete pricing:

Corollary 2 (Modified Square Rule and the Optimal Split Ratio). *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 split ratio to achieve the optimal price is $H^* = \sqrt{\frac{\mathbb{E}(s_t^{tot})-\Delta}{\Delta}}$.*

We test Corollary 2 empirically in Section 7. Before we move to the empirical tests, we analyze the impact of the regulator's choice in Stage 1 of the game.

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 tick and lot sizes are functions of price.

5.1. Tick and Lot Size Changes under a Uniform System

Under uniform tick and lot sizes, Corollary 3 specifies the Square Root Rule, under which the firm's best response, the change in normal price, and the change in liquidity all follow the square root of the change in tick and lot sizes.

Corollary 3. (Square Root Rule) *A firm's optimal nominal price* $p^* = \sqrt{\frac{\lambda_I \Delta v}{2\sigma\lambda_J L}}$ *responds to tick- and lot-size changes by* $\sqrt{\Delta/L}$. *When* $h \gg L$ *, the average nominal spread under optimal pricing equals* 2Δ *regardless of* L *and* Δ. *The smallest achievable execution cost* $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot Dvol\right) \approx \sqrt{2\sigma v \lambda_I \lambda_J \Delta L}$ *is proportional to* $\sqrt{\Delta L}$.

The comparative statistics of p^* show that a firm's optimal response to a change in tick or lot size is found in their square roots. For example, if regulators increase the tick size from one cent to five cents, firms should reverse-split their stocks by $\sqrt{5}$.[10] This reverse-split ratio is optimal because it changes 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.

Second, 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 ($= \sqrt{5} \times \sqrt{5}$)

[10] Here we consider a permanent fivefold tick size increase. The U.S. Tick Pilot program increased the tick size temporarily for two years. It is possible 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.

cents, and 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 response changes in accordance with the Square Root Rule.

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.[11] Corollary 3 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.[12] 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.

Corollary 3 shows that a policy initiative that aims to make prices (quantities) more discrete also makes quantities (prices) more discrete in equilibrium. To the best of our knowledge, we are the first to identify this spillover effect. Angel (1997) considers only discrete pricing, and he argues that 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 3 shows

[11] 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.

[12] Retail broker-dealers such as TD Ameritrade oppose the idea of reducing lot size. Their justification is that "display of unprotected quotes will cause confusion and mistrust in the market." Their comment letters can be found at https://www.ctaplan.com/oddlots.

that a 1-for-5 reverse split leads to the same transaction costs as doing nothing at all. The intuition behind this is as follows. 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. In equilibrium, 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 does. For example, consider a firm that currently has an optimal spread of two cents; one cent comes from the tick size and one cent comes 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 increase in the tick size is equivalent to doing nothing at all.

5.2 Proportional vs. Uniform Tick and Lot Sizes

One plan for changing the lot size is to make it a function of price, such that high-priced stocks have smaller lot sizes.[13] This plan essentially generates a proportional lot size, which leads to a uniform dollar lot size for all stocks. Also, in many European countries, Hong Kong, and Japan, the tick size increases with stock prices. Corollary 4 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 4. (Proportional Tick and Lot Systems) *(1) With fixed* Δ *and proportional*

[13] 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."

lot size $\mathbb{L}(p) = k^L/p$, *where* k^L *is a constant, the firm's optimal choice is* $p^* \to \infty$ *and* $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot Dvol\right) = \sigma\lambda_J k^L$.*(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* $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot Dvol\right) = \frac{k^\Delta \lambda_I v}{2}$. *(3) With proportional tick* $\Delta(p) = k^\Delta p$ *and lot* $\mathbb{L}(p) = k^L/p, \mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot Dvol\right) \equiv \sigma\lambda_J k^L + \frac{k^\Delta \lambda_I v}{2}$ *for any* p. *Adopting the proportional system in (3) 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 4. The intuitions are as follows. Suppose the regulator chooses lot size $\mathbb{L}(p)$ and tick size $\Delta(p)$, where both can be constants. 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 the 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	$p^* = \sqrt{\frac{\lambda_I \Delta v}{2\sigma\lambda_J L}}$	$p^* \to \infty$
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 tick size is proportional but 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.

The most interesting contrast appears at the diagonal, where we compare uniform tick and lot sizes with proportional tick and lot sizes. Such a comparison depends on k^{Δ} and k^{L}. 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 4 shows that such proportional systems would 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 chooses to retain the relative tick size and dollar lot sizes for a \$30 stock. The proportional system would impose a tenfold wider tick size and a 0.1-fold larger lot size on a \$300 stock. If the \$300 stock was at its equilibrium with a two-cent bid–ask spread, its tick-driven spread increases to ten cents, and its lot-driven spread reduces to 0.1 cents, leading to an increase in the total spread from two cents to 10.1 cents.[14] 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 \$3 stocks at the same magnitude. Corollary 4 implies that, if regulators want to

[14] 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, firm increases both its dollar lot size and the price, leading to a quadratic relationship.

switch from a uniform system to a proportional system, they should not use any existing stock as the benchmark.

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 actually 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. higher dollar lot sizes and lower relative tick sizes), because the main friction comes from discrete pricing. Less liquid stocks endogenously choose lower prices (i.e. lower dollar lot sizes and higher 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.

As both k^{Δ} and k^{L} can equal 0, Table 1 also provides the results for continuous tick and lot sizes. As we focus in our paper on modeling 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), who propose a fully continuous exchange. The uniform tick and lot sizes offer one degree of freedom for firms to balance between discrete pricing and quantities by choosing the nominal price. The 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. If the regulator uses an existing stock as the benchmark, the regulator may harm the liquidity of other stocks. The biggest victims would be those stocks whose optimal nominal prices (and implicitly the stock characteristics) differ to the greatest extent from the benchmark stock.

6. CROSS-SECTIONAL TESTS

In this section, we test our model in the cross-section. In Section 6.1, we show that the cross-sectional variation in the bid–ask spread follows the Modified Square Rule: our three-factor model of liquidity can explain 81% of the cross-sectional variation in the bid–ask spread. In Section 6.2 we show that our model also explains 57% of the cross-sectional variation in the nominal price with only two variables.

6.1 A three-factor empirical model of liquidity

Corollary (1) implies a three-factor model of cross-sectional variation in the bid–ask spread. Supposing that $s_t^{tot} - s_t^{\Delta} = s_t^L$, we have

$$s_t^{tot} - s_t^{\Delta} = s_t^L = \frac{2\sigma\lambda_J L}{DVol_t} p_t^2. \qquad (13)$$

Taking the natural log on both sides, we obtain

$$log(s_t^{tot} - s_t^{\Delta}) = 2log(p_t) - log(DVol_t) + log(\sigma\lambda_J) + const. \qquad (14)$$

Following the literature, our variable of interest is the time-weighted average spread. As our horizon is one month or more, it is safe to assume that p_t has evolved over a long enough period of time such that the average widening effect $E(s_t^{\Delta})$ becomes Δ following Proposition 2. Therefore, we can write (14) in the form of an OLS test:

$$log\left(\overline{Spread} - \Delta\right)_i = \delta \cdot log(Price)_i + log(Volume)_i + log(Volatility)_i + \varepsilon_i. \qquad (15)$$

We use daily Trade and Quote (TAQ) data to compute the time-weighted bid–ask spread, trading volume, and the number of trades. We use Center for Research in Security Prices (CRSP) data to compute daily volatility in a given year. We also use CRSP to compute the daily average price and market cap. Our Modified Square Rule predicts that $\delta = 2$. The null hypothesis is $\delta = 1$: when the lot size does not impose a binding constraint on the bid–ask spread, $\overline{Bidaskspread} - \Delta$ should increase one-to-one in price.[15]

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.[16] Our main sample period is the year 2020, the most recent period for which we have one full year of data, and we conduct robustness checks using previous years. We winsorize our variables at the 1% level. Table

[15] Consider two stocks that are identical except their nominal prices. If lot size does not impose any constraint on the 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 percentage spread is the same for both stocks. Then, we'll have $\delta = 1$ to adjust for their mechanical differences in share price.

[16] 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).

2 presents the summary statistics of our sample.

TABLE 2
Summary Statistics

	Mean	Min	Q1	Median	Q3	Max	Std.Dev	N
Nominal Price ($)	43.3672	1.0021	7.6600	17.6400	43.9550	2884.5000	100.7182	3745
Bid-Ask Spread (cents)	16.8440	1.0000	2.8154	6.4811	16.1344	318.0653	32.4711	3745
Shares Outstanding (Million)	127.1963	0.5660	18.1500	40.0760	93.0230	8747.0920	416.0766	3745
Log(Number of Trades)	7.8902	2.3979	6.7699	8.1101	9.1225	12.6363	1.7740	3615
Log(Dollar Volume)	22.1884	12.2625	20.2485	22.3209	24.1133	30.1408	2.6402	3745
Log(Market Cap)	20.4515	14.6745	18.9261	20.3376	21.8844	28.0539	2.1292	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 of 2020 as our sample, and we take the annual average 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.

Panel A of Table 3 strongly rejects 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-factor 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-factor model of liquidity with two canonical benchmarks: Madhavan (2000) and Stoll (2000). The results reported in column (1) show that the R^2 of the three-factor 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-factor model includes only a subset of variables that have been included in previous benchmarks. This improvement in goodness of fit is surprising, because adding more explanatory variables should, at a minimum, mechanically increase the R^2. The results displayed in columns (4)–(8) provide two explanations of this surprising outperformance, as we elaborate below.

TABLE 3
Lot-driven Spread and the Square Rule

Panel A: Three-Factor Model of Liquidity

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	$Log(s_t^L) = Log(s_t^{tot} - \Delta)$					
Sample Period	2020	2019	2018	2017	2016	2015
$Log(Price_t)$	**2.09*****	**2.08*****	**2.12*****	**2.06*****	**2.08*****	**2.06*****
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$Log(Volatility_t)$	**0.96*****	**1.19*****	**1.16*****	**1.07*****	**1.17*****	**1.14*****
	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$Log(Volume_t)$	**-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-Factor Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable	$Log(s_t^L)$	$\frac{s_t^{tot}}{2}$(bps)	$\frac{s_t^{tot}}{2}$(bps)	$Log(s_t^L)$	$Log(s_t^L)$	$Log(s_t^L)$	$Log(s_t^L)$
Sample Period	2020	2020	2020	2020	2020	2020	2020
$Log(Price_t)$	**2.09*****		**6.20*****	**2.24*****	**2.32*****	**2.26*****	**2.38*****
	(0.03)		(1.56)	(0.03)	(0.04)	(0.03)	(0.03)
$Log(Volatility_t)$	**0.96*****					**0.68*****	0.03
	(0.05)					(0.06)	(0.06)
$Log(Volume_t)$	**-0.84*****	**-30.46*****	**-19.17*****	**-0.57*****	**-0.75*****	**-0.60*****	
	(0.02)	(1.49)	(2.67)	(0.03)	(0.07)	(0.03)	
$Log(MKTCAP_t)$		**19.08*****	**9.69*****	**-0.46*****	**-0.42*****	**-0.41*****	**-1.22*****
		(1.71)	(1.21)	(0.04)	(0.03)	(0.04)	(0.02)
$Log(Turnover_t)$							
$Log(\#Trades_t)$			**-7.12*****		**0.10*****		
			(2.72)		(0.07)		
*$Volatility_t$ * 10^2*		**5.40*****		**0.09*****			
		(0.53)		(0.01)			
*$Variance_t$ * 10^4*			**0.19*****		**0.01*****		
			(0.02)		(0.00)		
$1/(Price_t)$		**-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 lot-driven nominal spreads on the log of nominal prices, controlling for log(Volatility) and log(Volume). We take a snapshot of the most recent five years of U.S. listed common stocks as our sample, and we take the annual average 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. 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 model's dependent variable and Madhavan's (2000) and Stoll's (2000) independent variables. 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.

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 imply a monotonic relationship between the price and the percentage spread. In reality (as seen in Figure 1), the relationship between price and liquidity is U-shaped. Therefore, these two canonical benchmarks may mis-specify the relationship between price and liquidity, at least for recent years. One indicator of the misspecification is the coefficient estimate of the price. For example, the Madhavan (2000) specification shows that an increase in the price or a decrease in the relative tick size increases the percentage spread, despite overwhelming evidence showing that exogenous decreases in tick size reduce the percentage spread (see Bessembinder 2003; Albuquerque, Song, and Yao 2020). 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. An increase in price, however, reduces the relative tick size but increases the dollar lot size. The trade-off between tick and lot sizes also rationalizes why Stoll's (2000) specification shows that the price does not correlate with the percentage spread, although the economic reasoning in Stoll (2000) suggests that the price should matter. Our model provides an interpretation that applies to this puzzle: the price matters for liquidity; it just does not matter in a monotonic way.

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 change the specifications in Madhavan (2000) and Stoll (2000) in only one respect: we replace the dependent variable in their regressions, the percentage spread (the bid–ask spread divided by the price), with the log of the percentage lot-driven spread ((bid–ask spread – 1 cent) divided by price). This one-cent change makes a huge difference. 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. Also, the coefficients for prices become statistically more significant.

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 (2) of Table 3 show that adding the market cap only marginally improves the explanatory power. Interestingly, Stoll (2000) documents a similar puzzle in his sample period: the market cap has very weak explanatory power for the percentage spread, and an increase in the market cap can *increase* the percentage bid–ask spread (Table 1, p. 1481).

Our model provides the intuition that explains why the market cap has almost no additional explanatory power for the percentage spread. Madhavan (2000) and Stoll (2000) also control for the dollar volume, and our model suggests that the market cap becomes a redundant variable after we control for the dollar volume. 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 amount in transaction costs and *ceteris paribus* quotes the same bid–ask spread.

The results reported in column (6) show that adding the market cap to our three-factor 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. Thus, dollar volume is a much stronger predictor of stock liquidity: 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.

6.2 A two-factor empirical model of nominal prices

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

$$log(p_t^*) = \frac{1}{2} log(\lambda_I v_t) - \frac{1}{2} log(\sigma\lambda_J) + const. \tag{16}$$

Rewriting (16) as a cross-sectional test gives us:

$$log(Price)_i = \frac{1}{2} log(Dollar\ Volume)_i - \frac{1}{2} log(Volatility)_i + \varepsilon_i. \tag{17}$$

In Table 4 we report the test results for nominal prices. In column (1) we show that the two-factor model of volatility and dollar volume captures 57% of the cross-sectional variation in stock prices. We find that an increase in volatility decreases the nominal price. This result is consistent with the puzzle raised by Baker, Greenwood, and Wurgler (2009), who find *"a somewhat unexpected result is the effect of volatility, which suggests that volatile firms have a greater, not lesser, propensity to manage prices downward."* Shue and Townsend (2021) provide a behavior-based interpretation of this puzzle. They hypothesize that investors think in part about stock-price changes in dollars rather than percentage units, leading to more extreme return responses to news by lower-priced stocks. Their interpretation focuses on the impact of price on volatility, whereas our interpretation focuses on the impact of volatility on price.

We can interpret this result using either lot constraints or tick constraints. Because an increase in volatility increases 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 a higher percentage spread and are less constrained by the tick size. Therefore, these stocks can achieve their optimal two-tick spreads at lower nominal prices.

We also find that the nominal price increases with dollar volume. As stocks that trade in higher volumes tend to be larger stocks, we provide an interpretation for the observations in Baker, Greenwood, and Wurgler (2009) and Weld et al. (2009) that large stocks choose higher prices. An increase in the dollar volume reduces the percentage spread and the tick-size constraints. Therefore, firms tend to choose higher prices to relieve tick-size constraints.

TABLE 4
Two-Factor Model of Nominal Prices

	(1)	(2)	(3)	(4)
Dependent Variable		$Log(Price)_t$		
Log(Volatility$_t$)	**-0.99*****			**-0.93*****
	(0.04)			(0.04)
Log(Volume$_t$)	**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-factor model on the cross-sectional nominal price 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. 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 with our baseline model, while for columns (2), (3), and (4) we report results obtained when testing the predictive power of volatility and industry fixed-effects as suggested by Weld et al. (2009). 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.

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). 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 found in Weld et al. (2009) is that firms in the same industry may be subject to similar volatility.

In summary, we find that our model fits qualitatively with cross-sectional variations in nominal prices. The fit is less perfect than 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, it is this imperfect fit that enables us to identify the impact of prices on the bid–ask spread. If all firms chose their prices following our model, log(price) would correlate almost perfectly with log(volatility) and log(dollar volume), leading to collinearity. There are two possible, albeit not mutually exclusive, interpretations of the less perfect fit of firms' behavior than of traders' behavior. First, our model omits some important drivers of the firm's choice but not of the traders' choices. Second, firms respond to market-structure frictions to a lesser extent than traders do. Therefore, firms may end up with suboptimal nominal prices. We cannot rule out the first interpretation, but we find empirical evidence consistent with the second interpretation through stock splits.

7 LIQUIDITY AND RETURNS AROUND STOCK SPLITS

In this section, we use stock splits as a laboratory in which to test the implications of our model. In Section 7.1, we describe our data and sample. In Subsection 7.2, we show that our model matches changes in the percentage spread after splits. In Subsection 7.3, we show that most splits are correct because they tend to increase liquidity. In Subsection 7.4, we find that model-predicted changes in the percentage spread can explain the cross-sectional variation in announcement returns on splits.

7.1 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.[17] 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

[17] The sample period begins in the month in which the millisecond TAQ data become available.

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 and announcement dates, split-adjusted stock returns, and market returns around declare dates as well as control variables. We use millisecond TAQ data to calculate the time-weighted quoted bid–ask spread and the quoted NBBO depth. 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 from 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 unique 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 supports our hypothesis that execution algorithms slice and dice their latent interests into smaller dollar sizes after stock splits. Also, 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.

TABLE 5
Summary Statistics

	Mean	Min	Q1	Median	Q3	Max	Std.Dev	N
Split Ratio	1.9122	1.2500	1.5000	2.0000	2.0000	50.0000	1.5192	1196
Pre-split price ($)	59.4892	3.4940	31.7909	47.5750	70.7538	3311.0000	105.0183	1196
Post-split price ($)	33.1518	2.1500	21.4525	30.0075	40.3000	440.6400	20.2794	1196
Market cap ($MM)	8.5407	0.0087	0.4226	1.5114	4.4431	2206.9111	68.6016	1196
Ex-ante spread (cents)	15.6910	1.0426	4.2007	7.0209	16.6008	294.5383	24.7741	1196
Ex-post spread (cents)	9.0609	1.0259	2.7923	4.3014	8.9847	103.2231	12.4447	1196
Predicted spread change (cents)	-9.9398	-290.6029	-9.2555	-3.9763	-2.1379	-0.0237	19.6756	1196
Predicted spread change (bps)	-15.2249	-497.0664	-12.6581	-3.2174	-0.9905	7.7606	34.8616	1196
Announcement CAR (%)	2.7307	-28.6696	-0.1252	1.7940	4.2501	68.8310	5.8791	1196
Ex-date CAR (%)	0.2957	-38.9020	-1.8954	-0.0102	2.0070	156.6785	6.4979	1196
Ex-ante trading volume ($MM)	39.6936	0.0145	1.1941	9.1207	35.2821	491.6339	80.1535	1196
Ex-post trading volume ($MM)	41.4456	0.0135	1.6153	10.4192	38.6120	486.6143	82.7543	1196
Pre-split number of trades (thousands)	3.3934	0.0030	0.2850	1.0330	2.7690	365.7730	15.7035	1141
Post-split number of trades (thousands)	5.9279	0.0020	0.4825	1.5620	4.3233	852.7215	34.7155	1163
Pre-split institutional holding (%)	57.9026	0.0000	32.1769	66.0525	84.4367	99.1116	30.9838	942
Post-split institutional holding (%)	58.0483	0.0000	32.8861	66.0563	84.1892	97.8309	30.1347	924
Pre-split number of holders (thousands)	14.9088	0.0010	0.4500	1.9500	7.1820	1234.0000	65.8571	753
Post-split number of holders (thousands)	16.4495	0.0010	0.4410	2.0445	7.7995	1426.0000	71.5956	758
Log(number of holders change ratio)	0.0286	-6.9312	-0.1022	-0.0114	0.2046	4.0761	0.8038	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 a split-announcement date. The announcement and ex-date CARs are cumulated announcement returns during dates [T-1, T+1], following Grinblatt, Masulis, and Titman (1984). Shareholder numbers are taken from the years immediately before and after stock-split announcements, and the logs of the changes are reported following Amihud, Mendelson, and Uno (1999). Other pre-split variables are measured in the 180-day-to-60-day window before split-announcement days and post-split variables are measured in the 60-day-to-180-day window after split-implementation days.

7.2. Spread Changes around Stock Splits

In this subsection we show that changes in percentage spreads fit closely with the Modified Square Rule. We measure the bid–ask spread before splits as the time-weighted average bid–ask spread 180 to 60 days before a split-announcement day.[18] We define R_i as the predicted change in the percentage spread:

$$R_i = \frac{(s_i^{pre} - \Delta)/H_i^2 + \Delta}{p_i^{post}/H_i} - \frac{s_i^{pre}}{p_i^{post}}, \tag{18}$$

[18] 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).

where $\frac{s_i^{pre}}{p_i^{post}}$ is the percentage spread before splits and $\frac{(s_i^{pre}-\Delta)/H_i^2+\Delta}{p_i^{post}/H_i}$ is the post-split percentage spread predicted by the Modified Square Rule.

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. 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 1 bp increase in the spread leads to a 1.02 bps realized increase in the spread, with *t*-statistics of 5.11. Therefore, the Modified Square Rule strongly predicts the percentage spread after splits. Also, after controlling for R_i, we find that the split ratio does not explain the change in the percentage spread.

TABLE 6

Predictions of Changes in Bid–Ask Spreads

Dependent Variable	Realized ΔS_i (bps)
R_i *(bps)*	**1.02*****
	(0.20)
Log(H_i)	0.17*
	(0.10)
Controls	Y
Industry-Year FE	Y
Obs.	1196
Adj. R^2	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 the 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 heteroskedasticities and within correlations clustered by firm. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

7.3. Correct versus Incorrect Splits

After showing that the changes in realized percentage spreads match almost one for one with the predictions derived from the Modified Square Rule, we consider whether splits improve liquidity. We find that firms in general make correct decisions for stock splits.

Our model predicts that a split is correct if it moves the bid–ask spread closer to the two-tick optimum. Mathematically, a split is correct if $R_i < 0$ and a split is incorrect if $R_i > 0$, and R_i reaches its minimum if the split ratio leads to the optimal two-tick spread. We find that 1,089 splits are "correct" and 107 splits are "incorrect." Among the 107 incorrect splits, 74 should have split, because their bid–ask spreads are higher 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 2-tick optimal. We find that R_i, on average, decreases by 15.22 bps in our sample, providing additional evidence that splits are in general correct.

7.4. Cumulative Abnormal Returns around Announcements

Liquidity affects asset value (Amihud and Mendelson, 1986), and Figure 4 presents preliminary evidence that our model-predicted liquidity change affects returns on split announcements: firms in the group with correct splits realize an average announcement CAR of 2.87%, whereas those in the group with incorrect splits obtain an average announcement return of only 1.36%.[19]

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

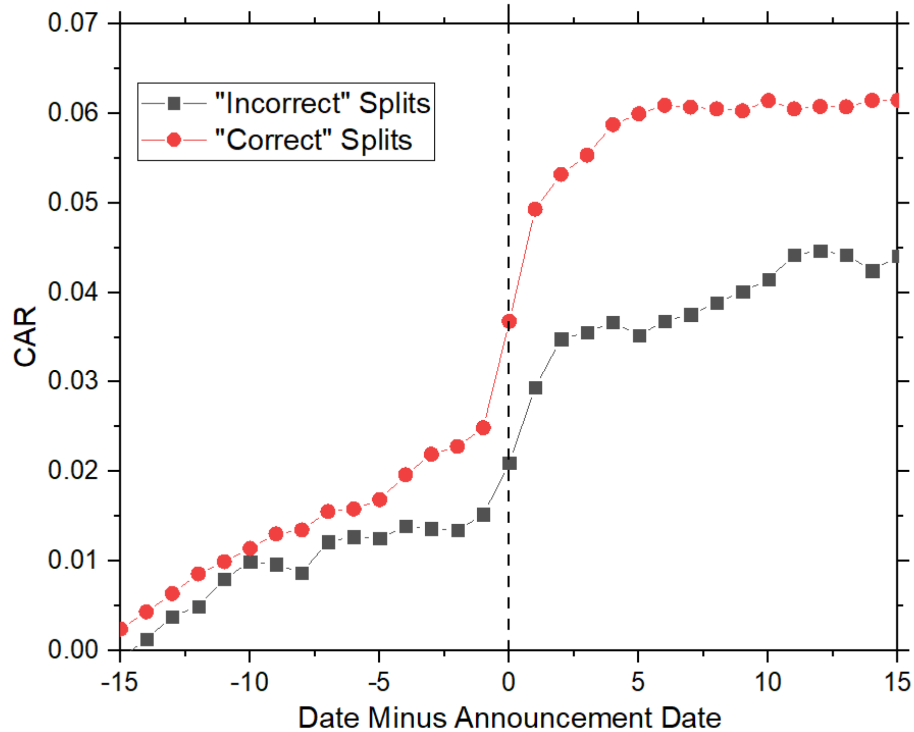

FIGURE 4.–Split-Announcement Returns: This figure shows the cumulative abnormal returns (CARs) around split-announcement dates. Our sample includes all U.S.-listed common stock splits beginning in September 2003. We require the firm to choose at least a $1 nominal price before and after a split. We categorize stocks into two types based on Proposition 3. A split is "correct" if Proposition 3 predicts a decrease in the percentage spread and "incorrect" if Proposition 3 predicts an increase in the percentage spread.

Insofar as splits are good news in general, 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:

$$CAR_{i,[T-1,T+1]} = \theta \cdot R_i + Controls_i + Industry_i \times Year\ FE_t + \varepsilon_i. \qquad (20)$$

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.[20] To reflect the information set of traders on the stock-split-announcement day, to avoid look-ahead bias we use the predicted spread change R_i but not the realized spread change.

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 1 bps increase in the percentage spread is associated with -5.47 bps in announcement returns.[21] After adding control variables, the results reported in column (3) show that a predicted 1 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 provides a partial explanation of 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. Brennan and Copeland (1988) predict, therefore, 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 a greater improvement in liquidity leads to a higher return. Both patterns are inconsistent with the signaling channel. 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 increased slightly after stock splits. The results reported in column (4) of Table 7 show that the change in retail holdings, proxied by the number of shareholders and

[20] 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.

[21] 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.

institutional holdings, does not affect announcement returns.

TABLE 7
Predicted Spread Changes and Abnormal Returns on Announcements

Dependent Variable	$CAR_{i,[T-1,\ T+1]}$ (bps)			
	1	2	3	4
R_i (bps)	**-5.47*****	**-4.42****	**-6.18*****	**-7.44****
	(1.48)	(2.06)	(2.40)	(3.72)
$Log(H_i)$		1.93	1.46	2.03
		(1.29)	(1.15)	(1.62)
$Log(MktCap_{i,t-1})$		**4.52*****	**4.81****	**7.93*****
		(1.59)	(1.80)	(2.28)
$Log(Price_{i,t-1})$		**-5.80*****	**-6.29*****	**-9.15*****
		(1.63)	(1.76)	(2.20)
$Turnover_{i,t-1}$		**5.53*****	**5.63*****	**9.16*****
		(1.59)	(1.81)	(2.20)
$Log(Volume_{i,t-1})$		**-4.87*****	**-4.96*****	**-8.08*****
		(1.54)	(1.72)	(2.17)
$Log(\frac{InstHldg_{t+1Q}}{InstHldg_{t-1Q}})$				7.38
				(3.73)
$Log(\frac{TOTSH_{t+1yr}}{TOTSH_{t-1yr}})$				-0.16
				(0.29)
Industry-Year FE	N	N	Y	N
Obs.	1196	1196	1196	607
Adj. R^2	0.067	0.132	0.164	0.237

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 (2) we control for log(market cap), price, log(volume), and turnover rates before the splits. Industry-year fixed effects are added to obtain the results reported in column (3) 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 (4) we also control for changes in institutional holdings and shareholders, following Dyl and Elliott (2006) and Amihud, Mendelson, and Uno (1999). 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 heteroskedasticities and within correlations clustered by firm. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

8. BACK-OF-THE-ENVELOPE CACULATION

In Sections 6 and 7, we provide evidence of the value of managing nominal prices, and we also show that firms on average move their nominal prices in the right direction during stock splits. Yet we still find many firms choosing suboptimal prices. The clearest evidence of this comes from pairs of stocks with similar fundamentals but dramatically different prices. For example, the daily average price is $3,305 for Amazon and $255 for Microsoft. Their dramatic differences in share price and their similar fundamentals indicate that at most one firm can be at the correct price. Which firm, then, is closer to its optimal price? How large is the benefit of moving to the optimal price?

The Two-Tick Rule indicates that Amazon is further away from its optimal price, which is consistent with the empirical evidence. Microsoft has a bid–ask spread of 1.95 cents, which is very close to the two-tick optimal. Amazon has a bid–ask spread of 153 cents, which indicates that its price is too high. Indeed, we find that Microsoft's percentage spread is 0.77 bps, which is much lower than Amazon's percentage spread (4.62 bps).

The sixfold difference in the percentage spread matches almost perfectly with the Modified Square Rule. Suppose that Amazon executes a 13-for-1 split such that its price is similar to the Microsoft price. The modified square rule predicts that Amazon's bid–ask spread should change to $\frac{153-1}{13^2}+1=1.90$ cents. The bid–ask spread becomes very close to Microsoft's bid–ask spread and the percentage spread experiences a sixfold reduction. The split ratio needed to generate the two-tick optimal spread is $\sqrt{\frac{1.53-0.01}{0.01}}=12.3$, and the percentage spread further reduces to 0.74 bps. The reduction in transaction costs would save Amazon investors $684 million per year. Amazon's market cap would increase by $3.89 billion based on our estimated elasticity of firm value to the percentage spread.

Figure 5 formalizes the intuition presented in the previous anecdote. We extend Figure 1 by adding three dashed lines that show the optimal percentage spreads if all firms in the basket choose the optimal prices predicted by our model, that is, nominal prices that generate two-tick bid–ask spreads. The horizontal axis presents the current price of the stock and the vertical axis presents the percentage spread. A larger vertical gap between the solid and dashed lines implies a greater gain in liquidity.

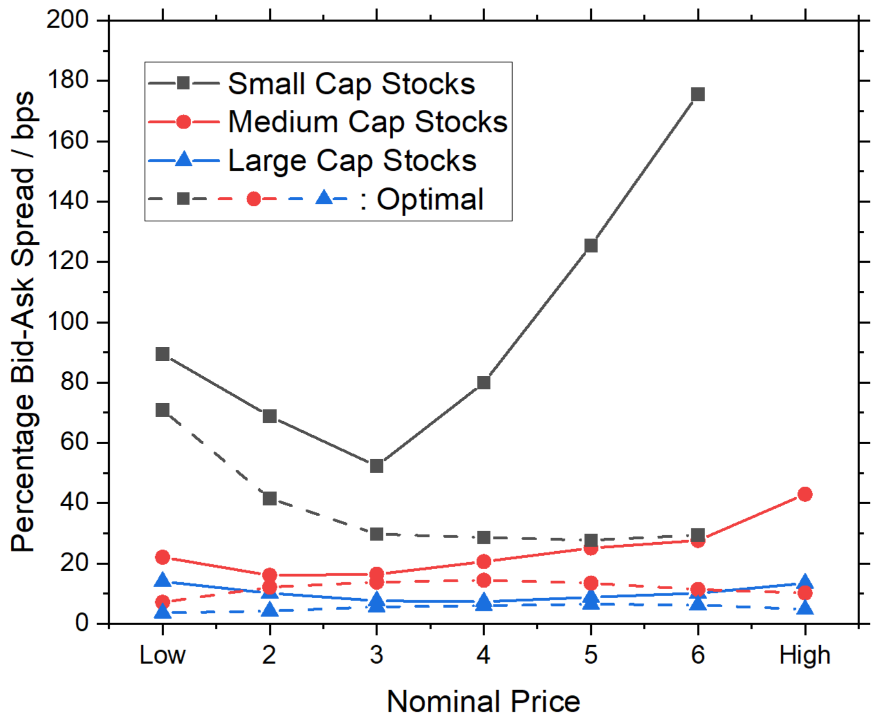

FIGURE 5.–Economic gains from adopting optimal nominal prices: This figure shows the relationship between average percentage spreads and nominal prices. Our sample includes all U.S.-listed common stocks that have a 1-cent tick size, a 100-share lot size, and at least a $1 nominal price. The lines with squares, circles, and triangles represent small-, medium-, and large-cap stocks, respectively. Price baskets are selected such that each basket contains a similar number of stocks. The solid lines are observed percentage spreads (simple averages) for stocks within the same price–size basket, and dashed lines are theoretically possible minimum spreads of the stocks in the same baskets.

For large stocks, the biggest winners would be those with low prices such as General Electric, Ford, Bank of America, and SiriusXM. Almost no firms voluntarily split in this segment of the market: most of these stocks experienced very large price slides before the sample period and have not fully recovered. Their low prices led to binding tick sizes, and they almost always trade at a one-cent spread, which results in a very large percentage spread. We conjecture that such firms do not reverse-split because reverse splits are usually regarded as negative "signals," and these firms would rather wait for (possible) price recovery to ease the binding tick size. We encourage these firms to ignore the negative connotations of reverse splits and escape the tick-binding restriction. One anecdote comes from competition between Ford and GM. Ford brought a price of $7 and GM brought a price of \$30. We then find that the Modified Square Rule can explain why Ford's percentage spread (14 bps) is four times greater than General Motors' (3.3 bps). If Ford were to implement a 1-for-4 reverse split, its percentage spread would be similar to General Motors'. Ford investors would save \$66 million per year if the company were to choose a price that is similar to GM's.

On the other side, a small firm should not choose a high price. Such a firm might choose a high price because people often consider a high-priced stock a prestigious stock (Weld et al. 2009). The cost of maintaining a minimum lot of 100 shares in liquidity is, however, very high for a small stock, leading to a large percentage spread for these stocks.

We find that the median optimal price for large stocks (NYSE deciles) is \$37.31, whereas the median small stock can sustain an optimal price of only \$3.51. The results presented in Figure 5 suggest that a small-cap stock should not choose a high price, whereas a large stock should not choose a low price.

Finally, we estimate the potential liquidity improvement that can be obtained with optimal pricing. After adopting optimal pricing, the median spreads will be reduced from 42.85 bps to 25.66 bps, a 40% reduction. For small-cap stocks, the spread will decrease from 153 bps to 91.7 bps. The spread for large-cap stocks will decrease from 14.00 bps to 7.91 bps. 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 $(42.85\text{-}25.66) \times 6.18 =$ 106 bps after adopting optimal pricing. Small stocks tend to be the biggest winners when achieving optimal nominal prices, and their value will increase by 378 bps if they choose their best nominal prices, but the median large-cap stocks can also increase their value by 37.6 bps. Summing up the potential gains for each stock, the total benefit of adopting optimal pricing is estimated to be \$93.7 billion. The top benefiting firms would be Alphabet Inc. (\$3.96 billion), Amazon (\$3.89 billion), and UnitedHealth Group (\$546 million).

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-factor model of liquidity outperforms two canonical 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 tick-driven spreads and then using log(price) as the independent variable to control for lot-driven spreads.

Our paper resolves several puzzles in behavioral finance. Baker, Greenwood, and Wurgler (2009) find it unexpected that volatile stocks have a greater propensity to manage their price downward. We can explain this puzzle by reference to lot size or tick size. An increase in volatility increases adverse-selection risk for market makers, and a firm should reduce its share price or dollar lot size to mitigate such risk. An increase in volatility increases percentage spreads, relieves the tick constraint, and gives a firm a stronger incentive to choose a lower price. 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 constraints. We find that volatility largely subsumes industry fixed effects. Therefore, firms 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. Firms have heterogenous optimal prices, but all firms reach their optimal prices when their bid–ask spreads equal two ticks, when tick-size friction is equal to lot-size friction. We find that most stock splits move bid–ask spreads closer to two ticks, and changes in liquidity following splits match almost one for one with the Modified Square rule. 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. We estimate that the median U.S. stock value 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.

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.

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.

Angel, J. J., Harris, L. E., & Spatt, C. S. (2015). Equity trading in the 21st century: An update. The Quarterly Journal of Finance, 5(01), 1550002.

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, 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.

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.

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.

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

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, forthcoming

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.

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. (2015). High frequency market microstructure. Journal of Financial Economics, 116(2), 257-270.

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.

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*

An uninformed trader demand h shares of liquidity, and she chooses the child order sizes to minimize her transaction costs. Suppose $f(q)$ is the frequency of child order of q lots. The uninformed investor chooses $f(q)$ for any q subject to $\sum_{q=1}^{h/L} qLf(q) = h$. We aim to show that $f(q) = 0$ for any $q > 1$ in equilibrium. That is, uninformed traders choose to slice their orders to minimum lots. We first solve the market maker's quoted spreads for any $f(\cdot)$ and then solve the uninformed trader's optimal $f(\cdot)$ through backward induction.

Observing the λ_I, λ_J, h, σ, and $f(\cdot)$, the competitive market maker quotes a liquidity schedule on both the bid and ask sides, and the number of layers of liquidity is $card\big(supp(f)\big)$. For example, if there exists (and exists only) X distinct order sizes $q_1 < q_2 < \cdots < q_X$ that satisfies $f(q_x) > 0$ for $\forall x \in \{1,2,\ldots,X\}$, the cardinality of the support of f is X, and the market maker quote X layers of liquidity on both sides of the market. The market maker supplies $q_1 L$ shares of liquidity in the first layer, $(q_2 - q_1)L$ shares in the second layer, and $(q_x - q_{x-1})L$ shares in the x^{th} layer. It's because for the x^{th} layer of liquidity, the market maker can trade only with liquidity demanding orders that are larger or equal to q_x round lots, which arrives at Poisson intensity $\sum_{i=x}^{X} \lambda_I f(q_i)$.

On the other hand, the market maker is also subject to adverse selection risks with Poisson intensity λ_J. Therefore, with probability $\frac{\sum_{i=x}^{X}\lambda_I f(q_i)}{\sum_{i=x}^{X}\lambda_I f(q_i)+\lambda_J}$, the market maker meets an investor and earns a spread of $\frac{s_t^x}{2}$. With probability $\frac{\lambda_J}{\sum_{i=x}^{X}\lambda_I f(q_i)+\lambda_J}$, a value jump occurs, the market maker is adversely selected, and the loss for the x^{th} layer is $\sigma p_t - \frac{s_t^x}{2}$. The equilibrium x^{th} layer of the bid-ask spread s_t^x should equalize the payoff from providing liquidity and sniping stale quotes, which leads to:

$$\frac{\sum_{i=x}^{X}\lambda_I f(q_i)}{\sum_{i=x}^{X}\lambda_I f(q_i)+\lambda_J}\frac{s_t^x}{2} = \frac{\lambda_J}{\sum_{i=x}^{X}\lambda_I f(q_i)+\lambda_J}\left(\sigma p_t - \frac{s_t^x}{2}\right).$$

$$s_t^x = \frac{2\sigma p_t \lambda_J}{\sum_{i=x}^{X}\lambda_I f(q_i)+\lambda_J}. \qquad \text{for } \forall x \in \{1,\ldots,X\}. \quad \text{(A.1)}$$

Therefore, $s_t^1 < s_t^2 < \cdots < s_t^X$, which indicates that the uninformed trader's large orders paid worse prices than her orders sized q_1. She can strictly reduce her execution costs by choosing $f(q) = \begin{cases} \frac{h}{q_1 L}, & for\ q = q_1 \\ 0, & otherwise \end{cases}$. [22] Furthermore, notice that $s_t^1 = \frac{2\sigma p_t \lambda_J}{\sum_{i=1}^{X} \lambda_I f(q_i) + \lambda_J}$ decreases in $\sum_{i=1}^{X} \lambda_I f(q_i)$. Since λ_I is a constant, the best f that maximizes $\sum_{i=1}^{X} f(q_i)$ is $f(q) = \begin{cases} \frac{h}{L}, & for\ q = 1 \\ 0, & otherwise \end{cases}$. In this case, we reach the one-layer minimum possible spread of

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

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 directly derive Corollary 1, which specifies that the dollar bid–ask spread $s_t^L = \frac{2\sigma \lambda_J L}{DVol_t} p_t^2$. ■

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 quotes that improve the bid and ask prices by one tick 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 distribution and the residual $mod(p_t, \Delta)$ tends to be 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 that v jumps up or down by σ% following a Poisson process

[22] Reducing all child orders to size q_1 reduces transaction costs through two channels. First, uninform traders no longer walk up the book and pay higher transaction costs s_t^2 ... s_t^X. Second, slicing orders to smaller size increases order arrive rate, and each child order would pay a cost lower than s_t^1

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 sufficient 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 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 in the order of $\frac{\Delta}{p}$. With this estimation, the residual of p within a tick is almost uniformly distributed.

Since $p \gg \Delta$, 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 in 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 $\frac{p}{\Delta}$ is smallest (i.e., when $p = \$1.005$ and

$f(\$1.00)/f(\$1.01) \approx 10^{-2}$ 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.[23]

We now solve s_t^{Δ} as a function of $mod(p_t, \Delta)$ and $s_t^L \equiv a\Delta + b$, where $a = 0,1,2,3, \ldots$ 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 $mod(p_t, \Delta) \in \left(\frac{\Delta}{2} - \frac{b}{2}, \frac{\Delta}{2} + \frac{b}{2}\right)$ (the "unlucky" case), the continuous pricing bid–ask spread can only be confined within $a + 2$ ticks. Otherwise, the "lucky" case arises.

The last step is to show that the widening effect is one tick with a uniformly distributed $mod(p_t, \Delta)$. 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 $\mathbb{E}(s_t^{\Delta}) = \frac{\Delta - b}{\Delta} \cdot (\Delta - b) + \frac{b}{\Delta} \cdot (2\Delta - b) = \Delta$.[24] ■

---

[23] In principle, any differentiable $f(p)$ with a bounded $f'(p)$ would lead to an approximately uniform distribution within a tick, as long as the variation in p is much larger than a tick so that in any neighborhood of a specific p, $f(p)$ does not exhibit large variation or a concentrated mass (Anshuman and Kalay 1985). 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.

[24] There is another intuitive way to understand this result. If the midpoint price p_t is uniformly distributed in the sub-tick granularity, the bid and ask prices $p_t \pm \frac{s_t^L}{2}$ are both uniformly distributed in the sub-tick

A.3. *Proof of Proposition 3*

Define h^* the shares outstanding under the optimal p^*, Equation (9) becomes

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

Recall Equation (4) that $s_t^L = \frac{2\sigma\lambda_J L p_t}{\lambda_I h + \lambda_J L}$. Inserting (A.6) into equation (4), the expected lot-driven spread under optimal pricing is

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

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

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

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

A.4. *Proof of Corollary 2*

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 be changed to $\frac{\mathbb{E}(s_t^{tot}) - \Delta}{H^2}$, and the tick-driven component remains Δ. Therefore, our theory predicts that the ex-post nominal spread is $\frac{\mathbb{E}(s_t^{tot}) - \Delta}{H^2} + \Delta$.[25] 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

granularity for any s_t^L. Therefore, the average widening effect on both the bid and the ask sides is $\frac{\Delta}{2}$, and the total widening effect is Δ.

[25] 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.

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

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, the 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 3*

From Proposition 3, we have $p^* = \sqrt{\frac{\lambda_I \Delta v}{2\sigma\lambda_J L}}$, which is proportional to $\sqrt{\frac{\Delta}{L}}$. Insert p^* into (7), we have

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

Note that $v = hp$, so the second term is negligible when $h \gg L$ and the expected transaction cost is proportional to $\sqrt{\Delta L}$. Thus, although the optimal nominal spread is 2Δ and does not depend on firm fundamentals, the optimal transaction cost does. Intuitively, more volatile firms ($\sigma\lambda_J$) need to choose lower prices to incentivize the market makers to quote the two-tick spread. On the other hand, firms with higher latent liquidity demand (λ_I) and larger market cap (v) can choose higher nominal prices to reach the two-tick nominal spread. ■

A.6. *Proof of Corollary 4*

In Equation (7), the expected transaction cost $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t} \cdot Dvol\right) = \sigma\lambda_J Lp + \frac{\Delta}{2}\lambda_I \frac{v}{p} + \frac{\Delta}{2}\lambda_J L$ depend on the firm's choice of p. Inserting the proportional lot size $L = \mathbb{L}(p) = k^L/p$, we have:

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

(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 is dependent only on k^L but not on p. In other words, the 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 (7), we have:

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

(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 $\mathbb{E}\left(\frac{s_t^{tot}}{2p_t}\cdot Dvol\right) = \frac{k^\Delta}{2}\lambda_I v$.

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

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

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 unified dollar lot size and relative tick size on all stocks. Next, we show that such a system harms liquidity provision if k^L and k^Δ is selected 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 tick (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 $p_t = p_\Omega$ (i.e., its tick and lot sizes are unchanged). The bid–ask spread widens for all stocks, with $p_t \neq p_\Omega$.■