Minimum Price Variations, Discrete Bid–Ask Spreads, and Quotation Sizes

Lawrence E. Harris
University of Southern California

Exchange minimum price variation regulations create discrete bid–ask spreads. If the minimum quotable spread exceeds the spread that otherwise would be quoted, spreads will be wide and the number of shares offered at the bid and ask may be large. A cross-sectional discrete spread model is estimated by using intraday stock quotation spread frequencies. The results are used to project $\$\frac{1}{16}$ spread usage frequencies given a $\$\frac{1}{16}$ tick. Projected changes in quotation sizes and in trade volumes are obtained from regression models. For stocks priced under \$10, the models predict spreads would decrease 38 percent, quotation sizes would decrease 16 percent, and daily volume would increase 34 percent.

Bid–ask spreads take discrete values. For most U.S. exchange-listed stocks, quoted spreads must be some multiple of the $\$\frac{1}{8}$ minimum price variation mandated by exchange rules.[1]

I am grateful to the editor (Chester Spatt) and referee (Larry Glosten) for their suggestions and to Jia Ye for her research assistance in the preparation of this article. Address correspondence to Lawrence Harris, Department of Finance and Business Economics, School of Business Administration, University of Southern California, Los Angeles, CA 90089-1421.

[1] Most exchanges require that quotes and transaction prices be stated as some multiple of a minimum price variation, or trading tick. The following rules are used in the primary U.S. stock markets. The minimum price variation on the New York Stock Exchange (Rule 62) is $\$\frac{1}{8}$ for stocks priced at and above \$1, $\$\frac{1}{16}$ for stocks under \$1 and at or above \$0.25, and $\frac{1}{32}$ dollar for stocks under \$0.25. The American Stock Exchange minimum price variation rule (Rule 127) was the same as the NYSE rule until September 2, 1992, when it was changed to sixteenths for stocks under \$5. The AMEX also allows some low-priced stocks formerly traded on NASDAQ to trade on sixteenths. The National Association of Securities Dealers permits trades on sixty-fourths

The Review of Financial Studies Spring 1994 Vol. 7, No. 1, pp. 149–178

The minimum price variation rules limit the minimum bid–ask spread that can be quoted. No quoted spread may be less than the mandated minimum price variation. Although this constraint usually is not binding for high-price stocks (especially if they are infrequently traded), it appears to be binding for low-price stocks and for some frequently traded stocks such as IBM. In 1989, 45 percent of all NYSE stock quotations had a spread of $\$\frac{1}{8}$.

The minimum spread constraint can be economically significant. The bid–ask spread for a \$2 stock is typically 6.25 percent of its price ($\frac{1}{8}$ divided by 2). The spread for an otherwise identical stock trading at \$40 with a $\$\frac{1}{4}$ quoted spread is 0.625 percent of price, or one-tenth as large. (The effective spread for such a stock is often lower since small market orders for stocks traded in a $\frac{1}{4}$ market often are executed at the quote midpoint.) Such spread differences surely affect trading decisions and may even affect stock valuations.

These observations suggest an obvious question: By how much would bid–ask spreads change if the tick were a different size? If the tick were decreased to $\$\frac{1}{16}$, many of the stocks quoted with a $\$\frac{1}{8}$ spread might be quoted with a $\$\frac{1}{16}$ spread. Some stock quotations might not change, and others might increase to a $\$\frac{3}{16}$ spread. Similar changes might be observed for stocks currently quoted at higher multiples of an eighth. This paper reports on an empirical analysis of this question for U.S. exchange-listed stocks.

The analysis also examines two related questions about how the minimum price variation affects displayed market depth and trading volumes. All three questions should be addressed when considering minimum price variation regulations.

The minimum price variation will affect displayed market depth (the sizes of the bid and ask quotations) when the minimum price variation is larger than the spread that dealers would otherwise quote. The spread then will equal the minimum price variation, and supplying liquidity could be quite profitable, especially to small orders. If the market enforces time precedence, dealers and other liquidity suppliers will queue up to offer liquidity and displayed depth will increase. If the market does not enforce time precedence, or if trade occurs in several markets that do not coordinate to enforce time precedence, dealers may offer inducements such as cash payments to obtain order flow.

A binding minimum price variation may also increase quotation sizes simply because dealers slide up their implicit quotation sched-

for all stocks. Quotes in the NASDAQ system, however, must be a multiple of $\$\frac{1}{8}$ if the bid is above \$10 and $\$\frac{1}{32}$ if the bid is under \$10. In practice, stocks below \$5 trade on sixteenths. These rules all permit exemptions. The NYSE and the NASD have not granted any exemptions for common stocks.

ules. The dealer quotation schedule is a schedule of prices for given quantities at which a dealer is willing to trade. As quantities increase, dealers typically require greater spreads to cover the risk of losing to large traders who are likely to be better informed than smaller traders.[2] At the major U.S. stock exchanges, dealers can only disseminate a single quote. If the minimum price variation is larger than the spread dealers would otherwise quote, they may choose to quote greater size.[3]

The minimum price variation may also affect displayed market depth even if the quoted spread is greater than the minimum price variation. The minimum price variation determines the minimum cost of acquiring order precedence through price priority when time precedence is enforced. Time precedence protects traders who expose their quotations and limit orders. By exposing their orders, these traders risk that other traders may act on this information to their disadvantage. In particular, some traders—call them quote matchers—may quote on the same side of the market when they see large size displayed. The quote matcher tries to profit on the information revealed by the large size, or to profit simply from the free trading option provided by the large size. In either event, quote matchers attempt to get their orders—orders that would not have been submitted had the large size not been displayed—filled ahead of the large size. Time precedence and a large minimum price variation protect traders who display size by forcing quote matchers to improve price significantly if they wish to acquire precedence. Displayed market depth should therefore be positively related to the minimum price variation.[4]

These arguments help explain why over-the-counter dealers typically quote small size. NASD markets, including NASDAQ, do not generally enforce time precedence among dealer quotations.[5]

---

[2] Easley and O'Hara (1987) provide a theoretical model of wider spreads based on the presumption that large traders are better informed. Glosten and Harris (1988) provide empirical evidence in support of the model.

[3] Ye and Harris (1994) provide a more complete analysis of how dealers quote a single point on the schedules when faced with a binding minimum price variation constraint.

[4] Amihud and Mendelsohn (1990) and Harris (1990b) provide detailed discussions of quote-matcher trading strategies and their effects on market liquidity. In particular, they note that the benefits, if any, of a time precedence rule will be difficult to obtain when time precedence is not enforced across markets.

[5] Time precedence is enforced among dealer quotes for dealers who participate in the Small Order Execution System. The SOES system, however, currently executes only about $1\frac{1}{2}$ percent of all volume.

At the New York Stock Exchange, Rule 2072 requires that time precedence be strictly enforced at a given price only for the first public bid (or offer) at that price. After a trade is made, time precedence ranks all remaining bids (or offers) that exceed any remaining size in the offer (or bid). If the offer (bid) was completely filled, all orders at the price are removed and reranked according to their resubmission times, or sizes if simultaneously resubmitted. After the first trade, floor orders tend to get precedence over the book because they are bigger. In practice, few stocks trade in a crowd so that time precedence is usually strictly enforced within the Exchange.

The minimum price variation will affect trading volume if it forces dealers to quote a larger spread than they would otherwise quote. Large bid–ask spreads make trading expensive, especially for smaller traders.

Most previous empirical studies of the cross-sectional determinants of the bid–ask spread do not model the effects of price discreteness on bid–ask spreads. These studies typically regress the relative spread (the quoted spread expressed as a fraction of price) on a number of cross-sectional variables, including price level, market value, volatility, trading volume, and trade frequency. The relative spread is used as the dependent variable because it measures transaction costs per dollar of investment. The price level is found to be a statistically significant determinant of the relative spread, even after controlling for these other variables.

These results would be perplexing if there were no price discreteness. Market microstructure models that ignore discreteness do not suggest that price level should be a determinant of the spread. Although price level is a proxy for recognized determinants such as firm size and trade frequency (and to a lesser extent, volatility), the inclusion of these variables in the regression presumably would lead to statistically insignificant price level coefficient estimates.[6]

Price discreteness can account for the statistically significant price level coefficient estimates. The minimum price variation may cause relative spreads for low-price stocks to be higher than they otherwise would be, considering only the standard market microstructure variables. Price determines the relative spread for such low-price stocks because price level determines the percentage size of the constant absolute minimum price variation. (The minimum price variation for almost all stocks that trade at prices above \$1 is fixed in absolute terms at \$$\frac{1}{8}$ at U.S. stock exchanges.) Price would also appear to determine relative spreads for higher-priced stocks, since price levels vary more than do absolute spreads.

This discussion suggests that average spreads may be better represented by a switching model. For low-price stocks (and some high-volume stocks) for which the minimum price variation is a binding constraint on absolute spreads, the inverse price level should be the only determinant of the relative spread. For high-price stocks and inactively traded stocks, most theoretical market microstructure models suggest that variables such as firm size, trading volume, and price

[6] Even if there were no price discreteness, statistically significant price level coefficient estimates could be obtained in a regression that includes size, volatility, and trade frequency as independent variables if the true regression is not linear in the (appropriately transformed) independent variables. Under these conditions, the correlation of price level with these independent variables could account for statistical significance in large samples. Statistical significance might also be due to the omission of some other variable with which the price level is correlated.

volatility should determine the proportional spread. These models do not, however, suggest that price level should be a significant spread determinant.

Most theoretical market microstructure analyses do not consider how the minimum price variation affects displayed size, as discussed above. If market depth and width are related, as many authors, including Kyle (1985) and Easley and O'Hara (1987) suggest, the minimum price variation may affect the bid-ask spread even for high-price stocks. In particular, most adverse selection and inventory models of the bid-ask spread suggest that dealers (and other liquidity supplying traders) implicitly maintain a schedule of bid and ask quotations for which spreads increase with size.[7] In practice, exchange dealers can only quote a single point on this schedule. Although few formal studies have been made of how they determine what point to quote,[8] the above discussion suggests that dealers may quote small spreads for small size if the minimum price variation provides inadequate protection against quote matchers. Price level may therefore be an indirect determinant of bid-ask spreads even for high-price stocks. A high price implies an economically small minimum price variation and therefore small displayed size and small bid-ask spreads.

This article presents several empirical analyses of spreads, quotation sizes, and trading volumes. Standard and switching regression analyses are used to explore how price levels (and indirectly the minimum price variation) are related in cross section to average spreads, quotation sizes, and trading volumes. The estimated regression models are then used to project how quotation sizes and trading volumes would change if traders could use a smaller minimum price variation. Finally, a discrete model of bid-ask spreads is introduced, estimated, and used to project how quoted spreads would change given a different minimum price variation.

The empirial results suggest that a decrease in the minimum price variation to $\$\frac{1}{16}$ would decrease bid-ask spreads by an average of 36 percent for stocks priced below \$10. Displayed size at the narrowed inside quotations for these stocks, however, would decrease by 15 percent. The net effect on liquidity would therefore be ambiguous. Small liquidity-demanding traders clearly would benefit from the change. Large liquidity-demanding traders would only be hurt if the total of displayed size at the narrower quotations, and size just behind the market where the quotations otherwise would have been placed, decreases. Otherwise, they also would be better off. The estimated

[7] Lee, Mucklow, and Ready (1993) provide some evidence of the empirical relation between spreads and depths.

[8] Ye and Harris (1994) is one such study.

models project that volume would increase by 30 percent for these low-price stocks. Similar qualitative effects would be observed for high-price stocks, but the quantitative effects would be small for all but the most actively traded stocks.

This article is related to several other studies about security-price discreteness. Gottlieb and Kalay (1985), Ball (1988), and Harris (1990a) show conditions under which stock-price discreteness may bias variance and serial covariance estimators. Ball, Torous, and Tschoegl (1985) and Harris (1990a) provide empirical evidence that suggests that price discreteness, price clustering, and price resolution are related.[9] The Harris study projects usage frquencies for odd sixteenth prices that would be observed if the minimum price variation for stocks were changed to $\frac{1}{16}$ of a dollar. The method characterizes fractional price usage frequencies for high-priced stocks for which the tick relative to price is small. The analysis then uses this information to project sixteenth-price usage frequencies for low-priced stocks for which the current tick is large relative to price. This study uses similar methods to project sixteenth-spread frequencies. Hausman, Lo, and MacKinlay (1992) introduce an ordered probit model for analyzing discrete stock-price phenomena. Their model characterizes discrete price changes, whereas the model in this article characterizes discrete bid–ask spreads.

The remainder of this article is organized as follows. Section 1 describes the data used in the study. Section 2 presents the results of some standard regression models and some switching regression models for spreads, quotation sizes, and trading volumes. Section 3 introduces a discrete model for bid–ask spreads. This model is estimated, and the results used to project spread usage frequencies that would be observed if the minimum price variation for stocks were changed to $\frac{1}{16}$ of a dollar. A short summary and a detailed discussion of the limitations of the results appear in Section 4.

1. Data

The data used in this study are obtained from the 1989 ISSM stock transaction and quotation data base. Only primary market (NYSE or AMEX) common stock quotations are analyzed. No stocks priced under \$1, no stocks that traded or were quoted on odd sixteenths, and no stocks that were quoted fewer than 150 times per calendar quarter are analyzed. The latter three restrictions eliminate approx-

[9] Price clustering in stock markets refers to the observation that whole numbers are more common than halves, halves are more common than odd quarters, and odd quarters are more common than odd eighths.

imately 27 percent of all stocks. The eliminated stocks are almost all infrequently traded. Stocks whose price levels changed by more than 50 percent in the sample period (mostly because they split) are also excluded from the sample. This restriction eliminated approximately 7 percent of all stocks.

Two samples of the data base are used. The two samples do not overlap in time or in cross section. The first sample contains data for approximately the first fourth of all NYSE and AMEX firms ranked in order by ticker symbol (A through EMR) for the first quarter of 1989. The second sample contains data for approximately the second fourth of the ticker-ranked firms (EN through MOT) for the second quarter of 1989.[10]

The models presented in this study were generally fully specified before estimation. In some cases, residual analyses of estimates obtained from the first sample suggested some minor changes in specification would produce better-fitting models. Although the specification searches were not extensive, Leamer (1978) and others show that specification searches can bias inferences made from model estimates. To determine whether the searches used in this study seriously bias the results, I reestimated the final specifications of each model, using the second sample. All model estimates are remarkably stable across the two samples, suggesting that the specification biases are not significant. The reestimation also shows that the results presented here appear to generalize across time and firms. Only results from the second sample are reported.

Cross-sectional summary means and correlations for the various variables of interest appear in Tables 1 and 2. The minimum price variation constraint on bid-ask spreads appears to be binding often. On average, 48 percent of the quotation spreads were $\$\frac{1}{8}$. Many of these spreads, of course, may be the result of rounding down to the nearest $\frac{1}{8}$ spread rather than rounding up. As expected, the $\frac{1}{8}$ spread frequency is higher for low-price stocks than for high-price stocks. It is not insignificant for high-price stocks, however: For stocks trading over \$40, one third of the quotations have a $\$\frac{1}{8}$ spread. These results suggest that the minimum price variation may be an empirically significant determinant of market quality for stocks of all price levels.

The high correlation between the inverse price level and the relative bid-ask spread is notable. The correlation is high because the variation in absolute spreads is small relative to the variation in price levels. Also note that the various measures of economic scale—firm size, transaction frequency, and dollar volume—are all correlated

[10] The exact composition of the two cross-sectional samples was determined by how much data could fit on a computer tape.

Table 1
Price level classified means of the regression variables

Variable	All stocks	Price level subsamples: Under \$10	\$10–\$20	\$20–\$40	Over \$40
OneEighthFreq	47.7%	66.8	46.9	34.6	33.3
RelSpread	1.76%	3.45	1.47	0.84	0.44
AvePrice	\$22.2	6.3	14.4	29.0	58.5
InvPrice	0.098	0.212	0.072	0.036	0.018
STDRet5	1.72%	2.29	1.65	1.38	1.25
LogMkVal	12.6	10.9	12.1	13.6	14.9
InvSqrtTrans	0.23	0.32	0.25	0.17	0.11
LogDolVol	4.76	3.80	4.60	5.30	6.11
LogPriMkShare	−0.191	−0.197	−0.182	−0.199	−0.178
N	529	163	146	145	75

The sample includes second quarter 1989 data from all 529 NYSE and AMEX securities with ticker symbols between EN and MOT that were priced above \$1, traded on eighths, and quoted more than 150 times. Only NYSE and AMEX data are used.

with each other. These correlations suggest that multicollinearity problems will be present in regression models in which all three variables appear.

2. Regression Models

This section presents cross-sectional regression models of average bid-ask spreads, average quotation sizes, and average trading volumes. The section concludes with a discussion of the simultaneous equations problems that arise when estimating these models.

2.1 Bid-ask spreads

Numerous studies examine regression models in which the average stock bid-ask spread, expressed as a fraction of price, is regressed on several explanatory variables. The explanatory variable set usually includes a measure of trading activity such as average daily transaction frequency and/or average daily volume, a measure of return volatility, perhaps a measure of the degree of information asymmetry such as firm size,[11] and perhaps a measure of how competitive is the dealer environment. Spreads are expected to decrease with trading activity because fixed costs of market making are spread over more traders, increase with volatility because dealers are risk averse and because volatility is probably correlated with information asymmetry, increase with the degree of information asymmetry because dealers must

[11] Firm size is a proxy for the degree of public information available about the stock. If the stock is well known, information asymmetries will tend to be small and the adverse selection component of the total spread should be small.

Table 2
Pearson correlation coefficient matrix for the regression variables

	One-Eighth-Freq	*Rel-Spread*	*AvePrice*	*InvPrice*	*STDRet5*	*Log-MkVal*	*InvSqrt-Trans*	*Log-DolVol*
RelSpread	0.43							
AvePrice	−0.48	−0.58						
InvPrice	0.52	0.98	−0.54					
STDRet5	0.10	0.54	−0.34	0.49				
LogMkVal	−0.20	−0.68	0.76	−0.61	−0.49			
InvSqrtTrans	−0.03	0.59	−0.51	0.49	0.28	−0.80		
LogDolVol	−0.52	−0.82	0.88	−0.78	−0.43	0.85	−0.65	
LogPriMkShare	−0.36	0.09	−0.04	−0.17	−0.01	−0.25	0.48	0.03

The sample includes second quarter 1989 data from all 529 NYSE and AMEX securities with ticker symbols between EN and MOT that were priced above $1, traded on eighths, and quoted more than 150 times. Only NYSE and AMEX data are used.

recover from uninformed traders what they lose to informed traders, and decrease with competition because competition eliminates monopoly power.

The set of explanatory variables also often includes the price level, although market microstructure models that ignore discreteness do not suggest that price level should determine spreads. Price discreteness should cause the estimated coefficient on price to be negative if multicollinearity problems are not large.

This typical regression analysis is replicated in this data set by using the following model specification:[12]

$$RelSpread_i = Desired_i + e_i,$$

$$\begin{aligned} Desired_i = {} & c_0 + c_1 AvePrice_i + c_2 STDRet5_i + c_3 LogMkVal_i \\ & + c_4 InvSqrtTrans_i + c_5 LogDolVol_i \\ & + c_6 LogPriMkShare_i, \end{aligned} \tag{1}$$

where

$RelSpread_i$ = average bid–ask spread computed over all primary market quotations, expressed as a percentage of price

$Desired_i$ = conditional mean relative spread ("Desired" reflects its interpretation in the switching model presented below)

$AvePrice_i$ = average price over the sample period

$STDRet5_i$ = five-day in-sample return standard deviation

[12] The slightly unusual notation in which the conditional mean of the regression is described by a separate equation is employed to simplify the presentation of the switching model that appears at the end of this subsection.

		computed from overlapping daily data (the five-day standard deviation is used to minimize the effects of discreteness on the volatility estimator for low-price stocks)
$LogMkVal_i$	=	log of the total average daily market value of the outstanding common stock equity
$InvSqrtTrans_i$	=	inverse of the square root of the average daily primary market transaction frequency[13]
$LogDolVol_i$	=	average daily primary market dollar volume
$LogPriMkShare_i$	=	ratio of average daily primary market volume to the average daily consolidated volume
e_i	=	independently, identically distributed error term

Ordinary least-squares estimates of this regression model appear in Table 3, column 1. The estimated variable coefficients for the volatility, transaction frequency, and dollar volume varibles have their expected signs and are significantly different from zero. The log primary market share variable coefficient estimate is negative. If large market share is a proxy for lack of competition among dealers, this result would be surprising. Note, however, that if spreads in the primary market are tight, order flow will migrate to the primary market and primary market share will be large, which is consistent with the observed result. The positive size variable coefficient and average price coefficients may both be due to multicollinearity. Both variables are correlated with the transaction frequency and dollar volume variables (Table 2).

The estimated coefficient for the average price variable is surprising because price discreteness should cause the estimated price coefficient to be negative. To demonstrate that multicollinearity explains the unexpected positive estimated price coefficient, I reestimated the regression after orthogonalizing the other explanatory variables with respect to the average price and intercept variables: Each explanatory variable was regressed on the average price and an intercept to obtain residuals that are orthogonal to average price. These residuals contain all information in these explanatory variables that cannot be explained by a linear function of price. The residuals are then used in the

[13] The transaction frequency variable appears in the model as a proxy for the profitability of market making in the stock. If there are few transactions, market makers will quote a large spread to cover their fixed costs, and they may be better able to extract surpluses in the absence of competition. These explanations suggest that the inverse number of transactions should appear. The square-root transformation is used because market makers in the public auction markets do not participate in all trades. Their participation rate declines with trading activity as more public orders cross. The transformation is also suggested by information-theoretic considerations: If each transaction conveys a bit of information about value, the standard error of the dealers's inferred estimate of value will be proportional to the inverse square root of the transaction frequency.

Table 3
Estimated bid-ask spread models

Variable coefficient	Model and estimation method					
	AvePrice OLS regression	OLS regression on orthogonalized regressors	*InvPrice* OLS regression	Nonlinear OLS switching regression	Instrumental 2-stage least-squares regression	MLE discrete bid-ask spread probability model
Intercept	7.82% (10.9)	2.91% (57.6)	−0.13% (−0.8)	4.35% (7.1)	−0.36% (−1.9)	0.55% (22.2)
AvePrice	0.048 (12.2)	−0.052 (−29.9)		−0.017 (−2.7)		
InvPrice			12.93 (96.2)		12.84 (62.4)	11.45 (153.9)
STDRet5	0.39 (10.1)	0.39 (10.1)	0.12 (11.7)	0.21 (7.0)	0.22 (7.3)	0.07 (26.4)
LogMkVal	0.15 (3.2)	0.15 (3.2)	0.01 (1.0)	0.02 (0.4)	0.02 (1.5)	−0.02 (−9.8)
InvSqrtTrans	2.9 (5.9)	2.9 (5.9)	1.8 (14.0)	2.1 (5.9)	1.6 (8.6)	1.3 (43.5)
LogDolVol	−2.24 (−20.7)	−2.24 (−20.7)	−0.04 (−1.5)	−0.81 (−7.0)	−0.04 (−1.1)	−0.05 (−12.9)
LogPriMkShare	−1.56 (−5.4)	−1.56 (−5.4)	−0.11 (−1.4)	−0.68 (−3.1)	0.08 (0.7)	−0.004 (−0.2)
Gamma shape parameter ν						6.8 (49.3)
R^2	0.806	0.806	0.987	N/A	0.981	N/A
RMSE	0.752%	0.752%	0.197%	0.229%	0.210%	N/A

The dependent variable of the regression models is the average relative bid-ask spread. The orthogonalized regressors are regression residuals obtained by regressing the explanatory variables on *AvePrice* and an intercept. The switching regression model takes a value of 0.125/*AvePrice* if the linear projection is less than 0.125/*AvePrice*. The instrumental variables estimation, described in Section 2.4, uses lagged values of all system regression variables as instruments. The discrete bid-ask spread probability model, described in Section 3, is a multinomial probability model estimated by using the vector of observed spread frequencies for each stock. The sample includes second quarter 1989 data from 529 NYSE and AMEX securities with ticker symbols between EN and MOT. t-statistics and asymptotic t-statistics are in parentheses.

regression model in place of the original explanatory variables. The resulting average price coefficient estimate is negative, as expected, and quite significantly different from zero (Table 3, column 2). The other explanatory variable estimates and t-statistics, the fitted regression values, and the regression residuals are all mathematically invariant to the orthogonalization. The negative price coefficient is consistent with a binding constraint placed on the absolute bid–ask spread by the minimum price variation for low-price stocks.

A plot of the estimated residual errors against the average price (not shown) shows strong J-shaped curvature. Such curvature suggests that a better-fitting model would be obtained if the inverse average price were used as an explanatory variable instead of the average price. This transformation, of course, is especially appropriate for low-price stocks for which the relative bid–ask spread is bound below by the minimum price variation.

Estimates of the inverse price model (Table 3, column 3) confirm that the fit is much tighter. The root-mean-squared error of the regression drops from 0.75 percent for the average price model to 0.20 percent for the inverse price model. The inverse price coefficient is now positive, as expected, and quite significantly different from zero. The log market value coefficient is now negative, as originally expected, but the log dollar volume coefficient is now positive. Both are insignificantly different from zero. The signs of the other coefficient estimates remain unchanged.

Note that the estimated inverse price coefficient (12.9) is greater than the minimum price variation, 12.5 cents. (The t-statistic for the difference is 2.98.) The difference between the estimated inverse price coefficient and the minimum price variation approximates the coefficient on inverse price that would be estimated if the dependent variable in this regression model were the excess of the relative spread over the relative minimum quotable spread. The latter is equal to the relative minimum price variation. Price thus appears to have some explanatory value in this regression beyond its effect on the relative minimum quotable spread. In particular, this result shows that high-priced stocks have narrower spreads, even after crudely accounting for the relative minimum quotable spread, and after accounting for the effects of the other regressors.

This positive difference between the estimated inverse price coefficient and the minimum price variation can be explained by the quote-matching argument presented in the introduction. This argument suggests that dealers will quote small spreads for small size if the minimum price variation provides inadequate protection against quote matchers. Since the minimum price variation is a larger fraction of price for low-price stocks than for high-price stocks, the protection it gives dealers against quote-matching strategies is greater for low-price stocks than for high-price stocks. High-price stocks should therefore have smaller spreads (and smaller quoted sizes) than low-price stocks, after controlling for the direct effect of discreteness on spreads and for all other microstructure factors that may cause spreads to vary in cross section. Alternatively, the positive difference between the estimated inverse price coefficient and the minimum price variation may simply be due to the omission of a correlated variable or to model misspecification.

These OLS model estimation results and the discussions in the introduction suggest that the average relative spread may be better represented by the following switching regression model:

$$RelSpread_i = Rounded_i + e_i$$

where

$$Rounded_i = \begin{cases} Desired_i & \text{if } Desired_i > \frac{1}{8} InvPrice_i, \\ \frac{1}{8} InvPrice_i & \text{otherwise,} \end{cases} \qquad (2)$$

and the high-price stock conditional mean, $Desired_i$, is given above in Equation (1).[14] Nonlinear ordinary least-squares estimates of this switching model appear in Table 3, column 4.[15] As expected, the switching regression model fits substantially better than the simple average price OLS regression model (RMSE of 0.229 percent versus 0.752 percent), although not quite as well as the simple inverse price model (RMSE of 0.197 percent). The average price coefficient is now negative. This result is analogous to the positive difference between the estimated inverse price coefficient and the minimum price variation observed in the previously reported regression. Both indicate that, after accounting for the relative minimum quotable spread and other cross-sectional differences among stocks, high-price stocks have narrower spreads relative to price than do low-price stocks. The signs

[14] If the desired relative spread for a given security varies through time, there may be stocks for which the minimum price variation is sometimes binding and sometimes not binding. This observation suggests that slightly better results can be obtained by using the following ad hoc specification:

$$Rounded_i = \begin{cases} Desired_i & \text{if } Desired_i > \phi \frac{1}{8} InvPrice_i \\ \phi \frac{1}{8} InvPrice & \text{otherwise,} \end{cases}$$

where ϕ is a parameter whose value is expected to be slightly greater than 1 to reflect the fact that the bound may not always be constraining. This specification was estimated, and the results (not shown) do indeed show some improvement in fit over those presented in the text. No qualitative differences, however, appear between the two sets of the parameter estimates. The simpler model is presented in the text because it is most similar to the standard OLS model presented above.

[15] The t-statistics reported for this regression are slightly misspecified since the regression model is discontinuous at the switch point.

of all other coefficient estimates are as expected, except for the log market value estimate. It is positive, but insignificantly different from zero.

The switching regression model estimates can be used to obtain a crude projection of the average spreads that would be quoted if the minimum price variation were reduced to $\$\frac{1}{16}$. The projection is computed by substituting $\frac{1}{16}$ for $\frac{1}{8}$ wherever $\frac{1}{8}$ appears in Equation (2). In addition, the average price coefficient is doubled. This coefficient adjustment is implied by market microstructure arguments: Suppose the average price coefficient is statistically significant only because it is a proxy for the inverse of the relative minimum price variation, $P/(\frac{1}{8})$. To project average spreads for a $\$\frac{1}{16}$ minimum price variation, this coefficient must be doubled to reflect the smaller tick. If the average price coefficient is statistically significant for reasons unrelated to the importance of the minimum tick, this adjustment would be inappropriate and the projected decrease in spreads due to the smaller tick would be overestimated.

The results suggest that spreads for the average stock in the sample would decline by 27 percent (8 percent if the average price coefficient is not adjusted). This projection may be poor since it is based on a crude model of how discreteness affects average quotation spreads. A detailed model of discrete quotation spreads is presented in Section 4. Projections from the discrete model probably are more reliable because the discrete process is better specified.

2.2 Quotation sizes

A change in the minimum price variation also may affect the sizes for which the best quoted prices are good. If dealers and other liquidity suppliers are concerned about quote matchers, they may choose to quote a lower point on their liquidity supply schedules for which both sizes and spreads are smaller. Alternatively, they may choose to quote a lower (or perhaps higher) point on their liquidity supply schedules simply because a smaller tick allows them to do so.

This subsection examines the relation between quotation sizes and price levels. Again, the inverse price level is interpreted as a proxy for the relative minimum price variation. Because price may be correlated with other variables that determine quotation sizes, the analysis explicitly includes such variables. If no missing variables are correlated with price and if the model is properly specified, the estimated coefficient for the inverse price level should provide information about the importance of the relative minimum tick.

The following regression model is used to describe average quotation sizes:

$$\begin{aligned} LogDolQSize_i = {} & c_0 + c_1 InvPrice_i + c_2 STDRet5_i + c_3 LogMkVal_i \\ & + c_4 InvSqrtTrans_i + c_5 LogDolVol_i \\ & + c_6 LogPriMkShare_i + c_7 RelSpread_i \\ & + c_8 OneEighthFreq_i + e_i. \end{aligned} \tag{3}$$

The dependent variable, $LogDolQSize_i$, the log of the average dollar quotation size, is the log average over all quotations for stock *i* of one-half of the sum of the bid times the bid size and the ask times the ask size. The log transformation is used to control heteroskedasticity across quotation sizes.

The inverse price is included in the regression to identify effects related to the minimum price variation. If liquidity suppliers fear quote matchers or if a large tick makes supplying liquidity profitable, the estimated inverse price level coefficient should be positive.

The return standard deviation measures price uncertainty. It should be negatively associated with quotation size because dealers are risk averse.

Log market value, transaction frequency, and dollar volume are all measures of firm size and trading activity. Each should have a positive effect on quotation size. (The inverse square root of the transaction frequency should have a negative effect on quotation size.)

Arguments concerning the sign of the primary market share variable lead to contradicting results: If market share is a proxy for competition, it should be negatively associated with quotation size, as competing dealers post size to solicit business. However, if quotation size significantly affects order flow, order flow may migrate to the dealer who posts the greatest size. The second argument is less convincing for size than it is for average spreads: It is unlikely that displayed size strongly affects order routing except for the largest of orders.

The relative bid–ask spread is included as a proxy for asymmetric information. Ye and Harris (1994) show that if the risk of trading with a well-informed trader is large, dealers will quote wide spreads and small sizes. In this cross-sectional regression, the spread should be negatively associated with size.

The percentage of quoted $\frac{1}{8}$ spreads, $OneEighthFreq_i$, is a proxy for whether the minimum price variation is a binding constraint on bid–ask spreads. Displayed size should be greater when this frequency is large.

The signs of the OLS regression coefficient estimates (Table 4, column 1) all appear as expected. The quoted $\frac{1}{8}$ spread frequency coefficient is of particular interest in this analysis. The positive and statistically significant estimated coefficient suggests that when the minimum price variation constraint on bid–ask spreads is binding, traders are willing to display more size. Traders probably are attracted

Table 4
Estimated OLS regression models for log average dollar quotation sizes and log average daily dollar volumes

Variable coefficient	Log average quotation size OLS	OLS	OLS	Instrumental 2SLS	Log average daily volume OLS	Instrumental 2SLS
Intercept	1.78	3.76	3.83	5.34	1.08	0.44
	(4.4)	(10.3)	(10.2)	(10.5)	(6.2)	(1.6)
InvPrice	1.44	9.82	2.17	3.29		
	(0.9)	(6.9)	(6.4)	(5.9)		
STDRet5	−0.16	−0.14	−0.22	−0.57	0.10	0.34
	(−6.0)	(−4.9)	(−8.2)	(−7.0)	(5.6)	(6.2)
LogMkVal	0.26	0.32	0.31	0.22	0.32	0.35
	(8.8)	(10.3)	(9.8)	(5.1)	(25.6)	(20.2)
InvSqrtTrans	−2.72	−3.01	−4.05	−4.25		
	(−8.0)	(−8.3)	(−12.8)	(−8.5)		
LogDolVol	0.45	0.09	0.11	0.15		
	(6.4)	(1.4)	(1.7)	(1.7)		
LogPriMkShare	−0.03	0.04	0.11	−0.05	0.93	1.12
	(−0.2)	(0.2)	(0.5)	(0.2)	(8.7)	(7.8)
RelSpread	−0.01	−0.59			−0.21	−0.28
	(−0.1)	(−5.5)			(−15.8)	(−12.8)
OneEighthFreq	0.015					
	(8.9)					
R^2	0.861	0.840	0.830	0.774	0.852	0.815
RMSE	0.449	0.481	0.495	0.566	0.341	0.386

The sample includes second quarter 1989 data from 529 NYSE and AMEX securities with ticker symbols between EN and MOT. The instrumental variables estimation, described in Section 2.4, uses lagged values of all system regression variables as instruments.

by high bid–ask spreads: Spreads presumably would be smaller if the minimum price variation constraint were not binding. They may also be displaying size that they would otherwise have displayed if they could quote inside the minimum spread.

The inverse price is proportional to the relative minimum price variation. Its estimated coefficient is positive, as expected, but it is not statistically significant. The statistical insignificance may be due to multicollinearity problems. The inverse price is correlated with the quoted $\frac{1}{8}$ spread frequency (Table 2). The correlation is primarily due to the low-priced stocks. When the quoted $\frac{1}{8}$ spread frequency is omitted from the regression (Table 4, column 2), the inverse price coefficient and the coefficient on the closely correlated relative spread variable both become statistically significant. To examine further how the multicollinearity problem affects the regression results, I reestimated the regression without the relative spread variable. The results (Table 4, column 3) are qualitatively unchanged, but as expected, the estimated inverse price coefficient is significantly smaller than in the second regression. The new estimate (which is measured in percent) approximates the second regression estimate plus $\frac{1}{8}$ of 100 times the second regression relative spread estimate. The other coefficient estimates are generally unchanged. The small differences in R^2 between the three regressions confirm that the multicollinearity problem is present in the first regression.

These results suggest that the minimum price variation may help determine quotation sizes. This conclusion is consistent with liquidity supplier fears about quote matching and with queuing to profit from spreads made wide (and profitable) by a binding minimum price variation, but it may simply be due to unrelated movements along the liquidity supply schedules.[16]

These three estimated regression models can be used to project the decrease in average quotation sizes that would occur if the minimum price variation were decreased to $\$\frac{1}{16}$. To use the first two equations, the effects of the change in the tick on the relative spread and on the quoted $\frac{1}{8}$ spread frequency must also be projected. These projections require the discrete bid–ask spread model that is examined in Section 3.

A simple projection can be obtained from the third regression equation in which only the inverse price level appears. Assume that the estimated inverse price coefficient is significant in this equation only because it is proportional to the relative minimum price variation. Halving the minimum price variation would then imply a halving of the estimated inverse price coefficient. The logarithmic regression for quote sizes therefore implies size will decrease by a factor of $\exp(-c_1 InvPrice_i/2)$ if the minimum price variation were decreased to $\frac{1}{16}$ from $\frac{1}{8}$ and if the change had no effect on the relative spread. Since the inverse price effect and the relative spread effect are approximately equally offsetting for low-price stocks, reliable projections can be obtained using the estimates from the third estimated regression model in which the relative spread is omitted as an explanatory variable (Table 4, column 3). The projected decrease in quotation size for a \$2 stock is 42 percent.[17] The projected decreases for \$5, \$10, \$40, and \$100 stocks are 20, 10, 2.7, and 1.1 percent respectively. The decrease would be greatest for low-price stocks for which the current minimum price variation is most economically significant.

2.3 Trading volumes

Any change in trading rules that affects bid–ask spreads will also affect trading volumes. Trading volumes depend, in part, on the bid–ask

---

[16] Note also that the negative estimated coefficient for the relative spread supports the asymmetric information hypothesis. This result complements the Lee, Mucklow, and Ready (1993) observation that spreads increase and sizes decrease before scheduled earnings announcements.

[17] This figure is computed as $100(1 - \exp\{-2.17(\frac{1}{2})(\frac{1}{2})\})$.

spread because the spread helps determine the cost of trading. Simple demand theory implies smaller spreads will be associated with larger volumes. Since market-maker profits are some rough function of the product of dollar volume times the relative spread, a change in trading rules might increase their profits if it increases volume by a greater percentage than it decreases the spread. This subsection examines a regression model for trading volumes to quantify the effect of the bid–ask spread on volume.

The following regression model describes average daily primary market dollar volumes:

$$LogDolVol_i = c_0 + c_1 STDRet5_i + c_2 LogMkVal_i + c_3 LogPriMkShare_i + c_4 RelSpread_i + e_i \quad (4)$$

where $LogDolVol_i$ is the log of the average daily dollar volume. The log transformation is used to control heteroskedasticity across volumes. The return standard deviation appears in the model because high volatility is said to attract traders. Log market value is the log dollar value of all outstanding shares. It helps scale the regression. Log primary market share is included to account for the fact that not all volume is primary market volume. In this logarithmic model, its estimated coefficient should be near 1. The estimated coefficient for the relative spread variable should be negative.

The signs of the OLS regression coefficient estimates (Table 4, column 4) all appear as expected. All estimates are significantly different from zero and the estimated coefficient for log primary market share is insignificantly different from 1 ($t = 0.69$).

The estimated relative spread coefficient of −0.21 in this logarithmic volume regression implies that log volume will increase by 21 percent for every one percentage point decrease in the relative spread. To obtain projections for the change in volume following a change in the minimum price variation, one needs a projection for the change in relative spread. This projection requires the discrete bid–ask spread model that is examined in Section 3. In the meantime, consider the following simple example: If the average bid–ask spread for a \$2 stock were to drop from $\$\frac{1}{8}$ to $\$\frac{1}{16}$ following a change in the minimum price variation, the projected increase in volume would be 94 percent.[18] This simple (and extreme) example suggests that market-maker profits for very low price stocks may not change much with a smaller minimum price variation: If the market maker's volume increases in proportion to the market volume, the projected near doubling of volume will almost offset the halving of the spread. More complete volume projections appear at the end of Section 3.

[18] The projected increase is computed as $100(\exp\{(0.21)(100)(\frac{1}{8} - \frac{1}{16})/2\} - 1)$.

2.4 Simultaneous equations and omitted variables

The various regression models that describe average bid–ask spreads, average quotation sizes, and average volumes suffer from the simultaneous equations problem. In particular, the bid–ask spread model includes volume as a regressor, the quotation size model includes spread and volume as regressors, and the volume model includes spread as a regressor. The use of endogenous variables as regressors may cause OLS estimation biases.

To examine whether these biases are significant, I estimated the system of three equations by using the instrumental two-stage least-squares method. The instrumental variables used were the previous quarter (first quarter 1989) values of the dependent and explanatory variables. The lagged values of the explanatory variables were used instead of the current explanatory variables to err on the side of caution: Some of the explanatory variables, such as transaction frequency, may also be endogenous. The results, reported alongside the OLS results in Tables 3 and 4, are similar to those obtained by using simple OLS.

These results suggest that the simultaneous equations bias does not significantly affect the OLS results in this study, but they do not rule out all specification biases. The OLS and 2SLS instrumental variable estimates can both have similar biases if important explanatory variables are omitted.

A simple example illustrates a possible omitted variable problem.[19] Suppose that spreads do not depend on volume, as specified in the above models. Instead, suppose that the average spread depends on whether the specialist is skilled. Skilled specialists have small spreads. The omission of a specialist quality variable can cause omitted variable biases in both the OLS and 2SLS estimations. If volume is determined by spreads, as specified above and suggested by demand theory, spreads and volume will be negatively correlated in cross section. This correlation will cause the volume coefficient in the OLS estimation to be negative. If the first-stage instruments are correlated with specialist quality, predicted volume obtained in the first stage of the 2SLS estimation will be correlated with specialist quality, because volume is inversely correlated with average spread, which is inversely correlated with specialist quality. The volume coefficient in the second stage also will be negative.

Specialist quality may also affect the results through the mechanism that allocates stocks to specialists. New stocks are allocated to specialists based in part on their skills. The more desirable stocks go to the more skilled specialists. These stocks are typically larger firms

[19] The author is indebted to the referee, who suggested this example.

that are more actively traded and that often have higher prices. Specialist quality therefore is correlated with many of the explanatory variables used in these analyses. These correlations may also bias the results. In particular, if skilled specialists quote smaller spreads, if high-priced stocks tend to be handled by skilled specialists, and if specialist quality is omitted from the analysis, the projected decrease in spreads for low-priced stocks from a change to a $\$\frac{1}{16}$ tick will be overestimated.

3. A Discrete Model for Bid–Ask Spreads

To more accurately project the spreads that would be quoted if traders could use sixteenths, one needs a discrete model for bid–ask spreads. This section specifies and estimates a model designed to represent the statistical properties of discrete spreads. The specification is ad hoc since maximizing behavior over discrete choices is not modeled. It is, however, sophisticated in its representation and organization of statistical information about discrete spreads, and much more so than the simple switching regression model of Section 2.1. This discrete bid–ask spreads model is analytically similar to the discrete data model used in Harris (1991) to analyze stock price clustering frequencies and to obtain projections for the sixteenth usage frequencies that would be observed if traders could use sixteenths.

3.1 Specification

The discrete quote generation process is represented by the rounding of random draws from a continuous distribution. The process for generating a quote at time t for stock i is as follows: First, an unobserved, underlying, unrounded relative spread, $RelSpread_{it}$, is drawn from a continuous distribution $F(RelSpread_{it}; m_t \nu)$ where m_t is the mean unrounded relative spread, ν is a distributional shape parameter, and F denotes the cumulative distribution function. The unrounded relative spread is then multiplied by the price level to obtain the unrounded absolute spread. The result is finally rounded to the nearest $\frac{1}{8}$ to obtain the observed absolute discrete spread, where nearest is defined relative to a logarithmic metric (geometric midpoints). The mean unrounded relative spread, m_t, is specified as in the regression models reported above for the average relative spread:

$$\begin{aligned} m_t = {} & c_0 + c_1 InvPrice_t + c_2 STDRet5_t + c_3 LogMkVal_t \\ & + c_4 InvSqrtTrans_t + c_5 LogDolVol_t + c_6 LogPriMkShare_t. \quad (5) \end{aligned}$$

The inverse price level is used to model the effect of the relative minimum price variation on the quotation spread, after accounting for discreteness.

The gamma distribution with degrees of freedom parameter ν is specified for the unrounded relative spread distribution. The gamma distribution ranges over all positive numbers, its shape parameter ν determines a rich family of distributional shapes, and it is computationally easy to use. The cumulative gamma distribution is

$$F(RelSpread_i;\ m_i,\ \nu) = \int_0^{RelSpread} \frac{\tau(\tau r)^{\nu-1} e^{-\tau r}}{\Gamma(\nu)}\, dr, \qquad (6)$$

where $\tau = \nu / m_i$.

The implications of this model for discrete spread frequencies are as follows:

$$\text{Prob}(Spread = \$\tfrac{1}{8}) = F(k_1/P), \qquad (7a)$$

$$\text{Prob}(Spread = \$\tfrac{1}{4}) = F(k_2/P) - F(k_1/P), \qquad (7b)$$

$$\text{Prob}(Spread = \$\tfrac{3}{8}) = F(k_3/P) - F(k_2/P), \qquad (7c)$$

$$\text{Prob}(Spread \geq \$\tfrac{1}{2}) = 1 - F(k_3/P), \qquad (7d)$$

where the geometric rounding midpoints are given by $k_i = [(i/8) \cdot (i + 1)/8)]^{½}$.

3.2 Estimation

The model is estimated with the maximum likelihood estimation method. The data consist of a vector of observed quotation spread size frequencies for each stock in the sample. The first through third elements of this vector, respectively, contain the observed frequencies for $\$\frac{1}{8}$, $\$\frac{1}{4}$, and $\$\frac{3}{8}$ quotation spreads. The fourth element contains the cumulative observed frequencies of all quoted spreads of $\$\frac{1}{2}$ or more.[20] The multinomial log likelihood for a given stock data vector is

$$\log L = T(\hat{\beta}_1 \log \beta_1 + \hat{\beta}_2 \log \beta_2 + \hat{\beta}_3 \log \beta_3 + \hat{\beta}_4 \log \beta_4), \qquad (8)$$

where T is the number of time-series observations, $\hat{\beta}_i$ is the observed frequency of spreads at $\$\frac{i}{8}$, and β_i is $\text{Prob}(Spread = \$\frac{i}{8})$ is given in Equations (7). For equal weighting of all stocks in the analysis, the time-series sample size T is taken to be the same for all stocks. So that the statistical inferences will err on the conservative side, T is

[20] Fewer than 5 percent of the quotes for stocks over \$40 are larger than $\$\frac{1}{2}$, and far less than 1 percent are larger for the lower-priced stocks.

Table 5
Cross-sectional means of projected average spreads and projected spread usage frequencies

Line variable	All stocks	Price level subsamples: Under \$10	\$10–\$20	\$20–\$40	Over \$40
1 Estimated underlying mean relative spread	1.64%	3.17	1.38	0.78	0.44
2 Observed average spread	20.7¢	16.8	20.4	23.5	24.2
3 Implied mean absolute spread	20.7¢	17.0	20.3	22.8	25.5
4 RMSError	2.7¢	1.4	2.5	3.1	3.9
5 Projected mean spread with $\frac{1}{4}$s	34.0¢	30.5	33.9	35.9	37.6
6 Projected mean spread with $\frac{1}{10}$s	18.5¢	14.8	18.5	20.6	22.3
7 Projected mean spread with $\frac{1}{16}$s	14.4¢	10.5	14.0	16.6	19.2
8 Projected mean spread with $\frac{1}{32}$s	11.2¢	7.4	10.9	13.5	16.1
9 Observed \$$\frac{1}{8}$ spread frequency	47.7%	66.8	46.9	34.6	33.3
10 Fitted \$$\frac{1}{8}$ spread frequency	47.8%	66.8	47.6	36.1	29.5
11 RMSError	10.8%	10.8	11.8	10.3	10.1
12 Projected \$$\frac{1}{4}$ spread frequency	67.0%	78.5	66.8	60.4	56.3
13 Projected \$$\frac{1}{10}$ spread frequency	37.7%	57.5	35.2	26.7	23.3
14 Projected \$$\frac{1}{16}$ spread frequency	24.2%	44.6	21.8	12.4	10.1
15 Projected \$$\frac{1}{32}$ spread frequency	8.7%	20.4	5.5	2.0	2.3
16 Number of stocks	529	163	146	145	75

These statistics are computed from the maximum likelihood estimates of the discrete bid–ask spread probability model described in Section 3 and presented in the last column of Table 3. This multinomial probability model is estimated by using the vector of observed spread frequencies for each stock. The sample includes second quarter 1989 data from 529 NYSE and AMEX securities with ticker symbols between EN and MOT.

assumed to be equal to 150, the minimum number of observations necessary for inclusion in the sample.

The signs of the estimated parameters (Table 3, last column) all appear as expected, and all the coefficients are significantly different from zero. The asymptotic t-statistics in this analysis are higher than the t-statistics obtained for the regression analyses. The difference arises because of the difference in models and because this discrete bid–ask spread model is a pooled time-series cross-sectional model.[21]

The cross-sectional mean of the estimated mean of the underlying relative spread is 1.64 percent for the entire sample (Table 5, line 1).[22] The mean estimate declines with price level: It is 3.17 percent for stocks priced under \$10 and 0.44 percent for stocks priced above

[21] If the time-series observations are truly independent, these asymptotic t-statistics understate the true significance because many stocks have more then the minimum 150 quotes assumed for each stock. If the time-series observations are not independent, the t-statistics will be overstated. If there were no independence in the time-series observations (for example, if all spreads for a stock were the same), corrected asymptotic t-statistics could be computed by assuming that the effective time-series sample length is 1. These can be computed by dividing the t-statistics in Table 3 by the square root of 150 ($1/\sqrt{150} \simeq 0.082$). Even under this extreme assumption, many coefficient estimates are significantly different from zero.

[22] Although the mean relative spread must be positive, the linear specification for it does not restrict it to be so. In practice, no negative means are estimated for any stock.

Table 6
Cross-sectional means of observed and fitted spread frequencies and their root-mean-squared differences for the estimated discrete bid-ask spread probability model

			Price level subsamples			
Spread size	Variable type	All stocks	Under \$10	\$10–\$20	\$20–\$40	Over \$40
$\$\frac{1}{8}$	Observed	47.7%	66.8	46.9	34.6	33.3
	Fitted	47.8	66.8	47.6	36.1	29.5
	RMSError	10.8	10.8	11.8	10.3	10.1
$\$\frac{1}{4}$	Observed	40.9	32.5	43.3	45.9	45.1
	Fitted	41.2	30.7	43.6	48.2	46.1
	RMSError	9.8	10.8	9.3	9.8	8.6
$\$\frac{3}{8}$	Observed	9.9	0.7	9.2	17.1	17.3
	Fitted	9.0	2.3	7.9	13.2	17.8
	RMSError	7.9	2.8	9.0	10.3	7.7
$\$\frac{1}{2}$ and over	Observed	1.4	0.0	0.5	2.5	4.3
	Fitted	1.9	0.1	1.0	2.6	6.5
	RMSError	4.3	0.2	2.7	5.1	7.9
	Number of stocks	529	163	146	145	75

These statistics are computed from the maximum likelihood estimates of the discrete bid-ask spread probability model described in Section 3 and presented in the last column of Table 3. This multinomial probability model is estimated by using the vector of observed spread frequencies for each stock. The sample includes second quarter 1989 data from 529 NYSE and AMEX securities with ticker symbols between EN and MOT.

\$40. The variation across price levels in almost every explanatory variable (Table 1) explains why the mean underlying relative spread declines with price level. Relative to low-price stocks, high-price stocks have low inverse price levels, low return volatilities, high market values, high transaction frequencies (and hence low inverse transaction frequencies), and high dollar volumes. The log market share variable does not vary significatnly across price levels.

The estimated model implies spread frequencies that closely fit the spread frequencies observed in the sample (Table 6). For the whole sample, the absolute difference between the mean implied (fitted) frequency and the mean observed frequency ranges between 0.1 and 0.9 for the various spread values. The fit is also good for stock subsamples classified by price level. Root-mean-squared errors of the fitted frequencies range between 0.2 and 11.8 percentage points across price levels and spread values. The fitted frequencies imply cross-sectional mean fitted spreads that closely fit the mean observed average spreads (Table 5, lines 2–4). The cross-sectional mean spread for all stocks is 20.7 cents; the mean average fitted spread is also 20.7 cents.[23]

[23] In regression analyses with constant intercepts, the cross-sectional mean observed value of the dependent variable and the mean fitted value must be equal. In this analysis, these variables need not be equal because this discrete data model is not a regression model and because the average spread is not the "dependent" variable.

Goodness of fit across the various frequencies is important because the estimated model is used to project $\$\frac{1}{16}$ spread usage frequencies. These projections are out of sample projections because the spread data on $\$\frac{1}{8}$, $\$\frac{1}{4}$, $\$\frac{3}{8}$, and $\$\frac{1}{2}$ spreads do not span $\$\frac{1}{16}$. The projections therefore crucially depend on the shape of the underlying relative spread distribution. The goodness-of-fit measures in Table 6 suggest that the gamma distribution usefully describes this distribution.

Goodness of fit across price levels is also important. The $\$\frac{1}{16}$ spread usage frequency projections for low-price stocks (and some very actively traded high-price stocks) depend on information about discrete spread usage learned from high-price stocks (and some low-price infrequently traded stocks) for which the current $\$\frac{1}{8}$ minimum price spread constraint is often not binding. These projections will not be reliable if the estimated model does not adequately describe both high- and low-price stocks.

The model specification assumes that traders always round to the nearest tick.[24] In practice, traders are unlikely to use any purely mechanical rule. Instead, traders may round conditional on the values of other variables.

To search for systematic misspecification in the rounding process, I computed the difference between the average spread and the fitted spread for each stock. This difference—the average rounding error—was then regressed on the explanatory variables used in the discrete spread model to determine whether the rounding is systematically related to the explanatory variables. (If the discrete spread model were a linear regression model, the rounding errors would be the residuals and the explanatory variables all would be orthogonal to them by construction.) The estimated regression coefficients and t-statistics are

$$\begin{aligned} RoundErr = & \underset{(-1.59)}{-4.02} - \underset{(-3.35)}{0.046}AvePrice + \underset{(1.32)}{0.0018}STDRet5 \\ & + \underset{(0.53)}{0.089}LogMkVal + \underset{(0.82)}{1.42}InvSqrtTrans_i \\ & + \underset{(1.72)}{0.66}LogDolVol_i - \underset{(-0.70)}{0.71}LogPriMkShare_i, \end{aligned} \qquad (9)$$

$$N = 529, \qquad R^2 = 0.032$$

Only the average price variable has statistically significant explanatory power. Apparently, the model underpredicts the spread for high-priced stocks. (This result can also be observed in lines 2 and 3 of

[24] If traders always round up (or down), the resulting model specification would be nearly isomorphic to the current model in which traders round to the nearest tick. All estimates will be the same except that the intercept would shift by approximately $(\frac{1}{16})P$, where P is the average stock price.

Table 5.) Similar but weaker results are obtained when the inverse price level is used as an explanatory variable. Although this analysis is quite crude and may exclude important variables, the low R^2 suggests that misspecification errors of this type are not empirically important.

3.3 Projections

The projected usage frequency of $\$\frac{1}{16}$ spreads given a minimum price variation of $\frac{1}{16}$ is computed from the estimated model by evaluating $\text{Prob}(\textit{Spread} = \$\frac{1}{16}) = F(k_{16}/P)$, where $k_{16} = [(\frac{1}{16})(\frac{1}{8})]^{1/2}$ and F is now the estimated gamma distribution. In addition, the estimated inverse price coefficient that appears in m_i is halved. This adjustment is analogous to the doubling of the average price coefficient in the projection obtained from the switching regression model. Both adjustments assume that the inverse price coefficient is statistically significant only because it is a proxy for the relative minimum price variation, $(\frac{1}{8})/P$. To project average spreads for a $\$\frac{1}{16}$ minimum price variation, the coefficient must be halved to reflect the smaller tick. If the inverse price is statistically significant in the model for reasons unrelated to the importance of the minimum tick, this adjustment would be inappropriate and the projected decrease in spreads due to the smaller tick would be overestimated.

The cross-sectional mean projected $\$\frac{1}{16}$ spread usage frequency for all stocks is 24 percent (Table 5, line 14). (The mean is 6.4 percent if the price coefficient is not adjusted.) As expected, the frequency declines with price level. The projected mean usage frequency for stocks trading under \$10 is 45 percent, but it is only 10 percent for stocks trading over \$40. (The corresponding mean percentages for the unadjusted price coefficient projections are 11.6 and 2.8 percent.)

Projected usage frequencies for all multiples of any minimum price variation are easily computed by similar methods. For example, the $\frac{3}{16}$ usage frequency can be projected by integrating the appropriate region under the underlying relative spread density.

Projected average spreads for a given minimum price variation are made by computing spread usage frequencies for all multiples of the minimum price variation. The projected average spread is then computed by the appropriately weighted sum of all possible spreads. The cross-sectional mean projected average spread for all stocks, given a minimum price variation of $\$\frac{1}{16}$ is 14.4 cents, which is 30 percent less than the current mean observed average spread of 20.7 cents (Table 5, lines 7 and 2). The reduction is mostly due to the adjustment of the estimated inverse price coefficient that shifts cumulative probabilities toward smaller spreads. Without the adjustment, the projected mean spread would be 20.0 cents, which represents only a 3.3 percent drop in average spreads. The percentage reduction in mean average spreads is greatest for stocks priced below \$10 (10.5 cents from 17.0 cents, a 38 percent reduction), while the reduction is 21 percent (19.2 cents from 24.2 cents) for stocks priced above \$40.

A matrix showing the complete projected mapping of spread frequencies from the various eighths to the various sixteenths can be constructed by mapping the factiles of the underlying spread distributions. The mapping is easy to do if the price coefficient is not adjusted. In that case, the underlying spread distribution is the same for all tick sizes. The map is constructed by identifying the corresponding rounded spreads for each possible underlying relative spread. When the price coefficient is adjusted, the underlying spread distributions differ by tick size. The mapping is accomplished by identifying the corresponding rounded spreads for each corresponding fractile of the two underlying relative spread distributions. The first method is a special case of the second method.

The projected spread frequency mapping (Table 7) shows that much of the projected decrease in spreads is due to the adjustment of the price coefficient. When the price coefficient is not adjusted, most of the decrease in spread is due to rounding more of the $\frac{1}{4}$ spreads to $\frac{3}{16}$ than to $\frac{5}{16}$. Interesting, more of the $\frac{1}{8}$ spreads are projected to be rounded up to $\frac{3}{16}$ than down to $\frac{1}{16}$.

Similar projections can be made for other minimum price variations. The projected $\$\frac{1}{32}$ spread usage frequencies, given a $\$\frac{1}{32}$ minimum price variation, are less reliable than the corresponding $\$\frac{1}{16}$ projections because they are further out of sample. The mean projected $\$\frac{1}{32}$ spread usage frequency is 20 percent for stocks trading under \$10 and only 2.3 percent for stocks trading over \$40 (Table 5, line 15). The corresponding mean projected average spreads are only 7.4 and 16.1 cents (Table 5, line 8).

Projections for a $\$\frac{1}{4}$ minimum price variation suggest that further tightening of the minimum price variation constraint could be very costly. The mean projected $\$\frac{1}{4}$ spread usage frequency is 78.5 percent for stocks trading under \$10 and 56.3 percent for stocks trading over \$40 (Table 5, line 12). The corresponding mean projected average spreads are 31 and 38 cents (Table 5, line 5).

Finally, the projections for a decimal $\$\frac{1}{10}$ minimum price variation suggest that this change to decimal pricing would not affect spreads much. The mean projected average spread is 14.8 cents for stocks trading under \$10 and 22.3 cents for stocks trading over \$40 (Table 5, line 6). These are 12 and 8 percent lower than the observed average spreads (Table 5, line 2).

The projected changes in average spreads can be used to compute projected changes in average volume from the OLS estimates of Equa-

Table 7
Cross-sectional mean projected transition frequency matrices for spreads based on eighths and sixteenths

Price level	Eighth spreads	Sixteenth spreads								
		$\frac{1}{16}$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{5}{16}$	$\frac{3}{8}$	$\frac{7}{16}$	$\frac{1}{2}+$	All

versity Salomon Center Monograph Series in Finance, 1990-4.

Harris, L. E., 1991, "Stock Price Clustering and Discreteness," *Review of Financial Studies*, 4, 389-415.

Hausman, J. A., A. W. Lo, and A. C. MacKinlay, 1992, "An Ordered Probit Analysis of Transaction Stock Prices," *Journal of Financial Economics*, 31, 319-379.

Kyle, A. S., 1985, "Continuous Trading and Insider Trading," *Econometrica*, 53, 1315-1335.

Leamer, E. E., 1978, *Specification Searches*, Wiley, New York.

Lee, C. M. C., B. Mucklow, and M. J. Ready, 1993, "Spreads, Depths and the Impact of Earnings Information: An Intraday Analysis," *Review of Financial Studies*, 6, 345-374.

Ye, J., and L. E. Harris, 1994, "The Specialist's Optimal Quotation Strategy," working paper, University of Southern California.