DECIMALIZATION, TRADING COSTS, AND INFORMATION TRANSMISSION BETWEEN ETFS AND INDEX FUTURES

ROBIN K. CHOU*
HUIMIN CHUNG

The impact of changes in trading costs, due to decimalization, on informed trading and speed of information transmission between exchange-traded funds (ETFs) and their corresponding index futures is examined. ETFs began to trade in decimals on January 29, 2001, and index futures continued to trade in their original tick sizes. The focus is on whether the decrease in the minimum tick size of ETFs influences the relative performances of these two types of index instruments in the price-discovery process. It is found that for ETFs, the trading activity increases, but the market depth drops significantly after decimalization. The spreads for ETFs generally decrease, but the adverse selection component of ETF spreads increases. Furthermore, after decimalization, ETFs start to lead

■ *Robin K. Chou is with the Department of Finance at National Central University in Taiwan.*

■ *Huimin Chung is with the Graduate Institute of Finance at National Chiao Tung University in Taiwan.*

WILEY
InterScience®
DISCOVER SOMETHING GREAT

index futures in the price-discovery process and its share of information also increases. Although index futures still assume a dominant role in information discovery, the information content of the ETFs' prices improves significantly after decimalization. © 2006 Wiley Periodicals, Inc. Jrl Fut Mark 26:131–151, 2006

INTRODUCTION

The recent move to a decimal system in quoting bid and ask prices at $0.01 increments in the U.S. equity markets has generated a great amount of interest. Both the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX) switched to the decimal pricing system on January 29, 2001. Exchange-traded funds (ETFs) listed on the AMEX started to trade in decimals on this date, but their corresponding index futures continued to trade in their original tick sizes. This offers an opportunity to test the impact of changes in tick size empirically.

ETFs are index funds or trusts that are listed and traded intraday on an exchange. In contrast to the traditional open-end mutual funds, ETFs allow for intraday trading and authorized participants can either create or redeem ETFs by delivering or receiving its index component stocks. In this study three actively traded ETFs are used as samples, which include the Standard & Poor's Depositary Receipts (S&P 500 ETFs), the Nasdaq-100 Index Tracking Stock (Nasdaq-100 ETFs), and the unit investment trust of the Dow Jones Industrial Average (DJIA ETFs).

The corresponding index futures contracts include the S&P 500 E-mini futures, the Nasdaq-100 E-mini futures, and the regular Dow Jones Industrial Average (DJIA) Futures.[1] Both the S&P 500 E-mini index futures and the Nasdaq-100 E-mini index futures are traded on GLOBEX, an electronic trading system operated by the Chicago Mercantile Exchange (CME). The DJIA futures contracts are traded on an open-outcry system of the Chicago Board of Trade (CBOT).

The relative changes in spreads, adverse selection, and information transmission between ETFs and their corresponding futures surrounding decimalization is examined. The study provides several valuable contributions to the existing literature. First, the impact of decimalization on the trading costs and liquidity of index instruments; that is, ETFs and index futures is examined. This is an area that receives relatively little attention in the literature. Bollen, Smith, and Whaley (2003) examine

---

[1]E-mini futures are more actively traded than the regular futures and thus are less influenced by the nonsynchronous trading problem. Because the E-mini version of DJIA futures started trading on May 1, 2002, it was not available surrounding the date of decimalization. Regular DJIA futures are used instead.

the effect of an increase in tick size of S&P 500 futures. They find increases in the S&P 500 futures bid–ask spreads, but the spreads remain low relative to those of S&P 500 ETFs, which is consistent with the empirical results here. DeJong and Donders (1998) examine the relations between futures, options, and index levels, and find that futures significantly lead options and index returns. It is known that index returns are likely to be affected by the nonsynchronous trading problem; thus ETFs, which are traded index instruments and are less likely to be affected by the nonsynchronous trading problem, are examined.

The second contribution is a test of which index instrument informed traders tend to exploit for their information advantage. Other things being equal, the potential profits of information trading are higher for futures because of their higher leverage effects (Kawaller, Koch, & Koch, 1987). Decimalization makes the minimum tick size of ETFs relatively smaller than that of the index futures. Informed traders now have more incentives to trade ETFs because of lowered costs. Beaulieu, Ebrahim, and Morgan (2003) study lead–lag relations between the Toronto Stock Exchange (TSE) 35 Index futures and the TSE 35 Index Participation Units before and after a decimal pricing system is implemented. They find that ETFs start to lead futures after their tick size is reduced. Similar results are found in this study, but in addition to testing of the lead–lag relations between futures and ETFs, the size of adverse selection components and information shares is also examined to provide additional evidence on the impact of decimalization on informed trading.

The third contribution is to test the impact of changes in trading costs on the price-discovery process. Fleming, Ostdiek, and Whaley (1996) argue that the relative rates of price discovery in the stock, futures, and options markets are due to differences in trading costs, which they refer to as the trading-cost hypothesis. They show that markets with lower trading costs tend to lead those with higher trading costs in price discovery. Decimalization of ETFs offers another opportunity to test the trading-cost hypothesis.

The empirical results indicate that, consistent with the literature on equity securities, both spreads and depth of ETFs decline significantly after decimalization, while trading activities of ETFs and index futures generally increase. The adverse selection component for ETFs increases, indicating that after decimalization, informed traders seem to trade ETFs more intensively. Furthermore, ETFs start to lead index futures in the price-discovery process and the information shares of ETFs also tend to increase after decimalization. Overall, decimalization improves the quality of the ETFs market, in terms of trading costs and information efficiencies.

The rest of this article is organized as follows. The data and research methodology are described. Empirical results are presented, followed by a conclusion.

DATA

The trade and quote prices of ETFs are retrieved from the Trade and Quote (TAQ) database published by the NYSE. The index futures trade prices are obtained from the Tick Data Inc. The sample period starts on October 29, 2000 and ends on April 28, 2001, which spans 3 months before and after the ETFs' decimalization date. The ETFs data include the tick-by-tick trade and quote prices, trade volume, and quote size behind the best bid and offer (BBO) prices.[2] The minimum tick size for ETFs is $1/16 before decimalization and $0.01 after decimalization. Throughout the entire sample period, the minimum tick sizes for S&P 500 E-mini and Nasdaq-100 E-mini futures contracts are 0.25 and 0.5 index points, respectively, and the minimum tick size for DJIA futures is 1 index point.

For index futures, only the tick-by-tick trade prices are available.[3] This poses a problem for measuring quoted spreads and trading activity of index futures. For index futures, two implicit spread measures are used instead in estimating the bid–ask spreads, and trading activity is measured by number of trades. The first implicit spread measure is the Roll's (1984) implied spread. The second implicit measure is the estimator suggested by Wang (1994) and is also used by the Commodity Futures Trading Commission (CFTC) in estimating futures spreads.[4]

RESEARCH METHODOLOGY

Trading costs, trading activity, information transmission, and information shares for ETFs and index futures are computed and compared. These measures are explained in the following sections.

Measures of Trading Costs

Spreads are common measures for trading costs. The relative quoted spread (QS) is calculated as

$$QS_{it} = \frac{(A_{it} - B_{it})}{M_{it}} \tag{1}$$

[2]Ideally, the quote size beyond the BBO needs to be tested before a definite conclusion about market depth can be drawn. Nevertheless, such information is generally unavailable from public databases.
[3]This is also the case for most other futures studies.
[4]All futures prices used in this study are those of the nearby contracts. The rollover of each contract is made on the ninth day before the last trading day to avoid any expiration effects.

where A_{it} (B_{it}) is the quoted ask (bid) price for stock i at time t, and M_{it} is the midpoint of the quoted ask and bid prices.

Relative quoted spreads are likely to be biased estimators of trading costs, because trades do not always occur at the posted quotes. The relative effective spread (ES) measures the difference between the actual traded price and the midpoint of the quoted bid and ask prices and provides a better measure of the actual trading costs. It is calculated as:

$$\mathrm{ES}_{it} = 2 \left(\frac{|P_{it} - M_{it}|}{M_{it}} \right) \tag{2}$$

where P_{it} is the transaction price for stock i at time t and M_{it} is the midpoint of the bid and ask prices of the quotes immediately prior to the transaction. The quote is required to be at least 5 seconds before the trade.[5]

Because quotes of index futures are not available, quoted and effective spreads can only be estimated for ETFs. Two more implicit spread measures are applied in estimating the spreads for index futures, as well as for ETFs. The first implicit spread measure is the Roll's (1984) implied spread. Roll suggests a simple procedure to estimate the effective spread based on the observed return covariance. However, as has been documented in the literature, it is not uncommon for the Roll's method to generate negative spread estimates.[6] A simple regression framework that does not require discarding negative spread estimates is applied.[7] The regression equation is:

$$r_t = \alpha + \tfrac{1}{2}s \cdot \Delta Q_t + u_t \tag{3}$$

where α is an intercept term and Q_t is a trade indicator that is $+1(-1)$ if the trade is buyer (seller) initiated.[8] By running this regression, it is not necessary to drop any negative spread estimates.

The second implicit spread measure is suggested by Wang (1994), and is also employed by CFTC. It is calculated as the average opposite

[5]Lee and Ready (1991) note that trades are often reported with a delay. They recommend using the quotes that are time stamped at least 5 seconds before the current trades.

[6]See Amihud and Mendelson (1987), Hasbrouck and Ho (1987), and Kaul and Nimalendran (1990) for evidence of a positive autocorrelation in short-term asset returns. George, Kaul, and Nimalendran (1991) suggest that a positive covariance may be due to partial price adjustments as dealers attempt to smooth price changes, or due to time-varying expected returns. Furthermore, the order-splitting strategy by institutional traders may also induce a positive autocorrelation in stock returns.

[7]This approach is similar in spirit to those used in Neal and Wheatley (1998), and Van Ness, Van Ness, and Warr (2001), in modifying the covariance methods of spread decompositions.

[8]When classifying trade signs, the Lee and Ready algorithm for ETFs and the tick rule for index futures are used, because the index futures' quote data are not available.

direction absolute price change. Price changes in the same direction as the preceding price change are discarded so as to reduce the impact of changes in the underlying futures price unrelated to the bid–ask bounce. This measure is to proxy for the average magnitude of the bid–ask spread.

Decomposition of the Bid–Ask Spreads

Gibson, Singh, and Yerramilli (2003) show that, for a sample of S&P 500 stocks, the percentage adverse selection component increases significantly after decimalization. They conclude that a smaller tick size increases traders' incentives to gather information. To test whether informed traders are more willing to trade ETFs, relative to index futures after decimalization, four methodologies—Glosten and Harris (1988) (hereafter, GH); George, Kaul, and Nimalendran (1991) (hereafter, GKN); Huang and Stoll (1996) (hereafter, HS), and Madhavan, Richardson, and Roomans (1997) (hereafter, MRR)—are used to decompose the adverse selection component of ETF spreads.

GH express the adverse selection component and the combined order-processing and inventory-holding component as linear functions of transaction volume. The model is

$$\Delta P_t = c_0 \Delta Q_t + c_1 \Delta Q_t V_t + z_0 Q_t + z_1 Q_t V_t + e_t \tag{4}$$

where P_t is the observed transaction price at time t, Q_t is the trade indicator signed by the Lee and Ready algorithm, V_t is the trading volume, and e_t is the residual term. The adverse selection component is $Z_0 = 2(z_0 + z_1 V_t)$ and the order-processing and inventory-holding component is $C_0 = 2(c_0 + c_1 V_t)$, which sum up to the bid–ask spread. An estimate for the proportional adverse selection component of the spread is

$$Z = \frac{2(z_0 + z_1 V)}{2(c_0 + c_1 V) + 2(z_0 + z_1 V)} \tag{5}$$

where V is the average trading volume.

GKN show that the difference between transaction returns and bid returns can filter out the serial dependence in returns. The resulting estimate of the adverse-selection component is expressed as:

$$2RD_t = \pi s_q \Delta Q_t + u_t \tag{6}$$

where RD_t is the difference between transaction return and bid-to-bid return immediately following the transaction return at time t, π is the order-processing component, $1 - \pi$ is the adverse selection component, s_q is the percentage quoted bid–ask spread, Q_t is the trade indicator

defined by the Lee and Ready algorithm, and u_t is the disturbance term. Adding an intercept to the above equation and relaxing the assumption that s_q is constant lead to:

$$2RD_t = \pi_0 + \pi_1 s_{qt} \Delta Q_t + u_t \tag{7}$$

Following HS, the after-trade price reversals are measured as the market-making revenue net of losses to better-informed traders, that is, the order-processing component. The adverse selection (AS) and order-processing components (OP) are calculated as

$$AS_{it} = Q_{it}(P_{it+n} - M_{it})/M_{it} \tag{8}$$

$$OP_{it} = Q_{it}(P_{it} - P_{it+n})/M_{it} \tag{9}$$

where Q_t is the trade indicator signed by the Lee and Ready algorithm and P_{it+n} is the first trade price at least 5 minutes after the trade at time t.

MRR show that u_t, the posttrade expected value from their model's price generation process, can be expressed as

$$u_t = p_t - p_{t-1} - (\phi + \theta)x_t + (\phi + \rho\theta)x_{t-1} \tag{10}$$

where p_t is the trade price at time t and x_t is a trade indicator. $x_t = 0$ if a trade takes place within the prevailing bid and ask prices. θ is the adverse selection component, ϕ is the cost of order processing, λ is the probability that a trade takes place inside the spread, and ρ is the autocorrelation of order flow. As in MMR, the generalized method of moments (GMM) is used to identify the parameters and a constant drift α implied by the model

$$E\begin{pmatrix} x_t x_{t-1} - x_t^2 \rho \\ |x_t| - (1 - \lambda) \\ u_t - \alpha \\ (u_t - \alpha)x_t \\ (u_t - \alpha)x_{t-1} \end{pmatrix} \tag{11}$$

To obtain the proportional adverse selection and order processing components, θ and ϕ are estimated in dollar terms and divided by the mean effective half spreads for each ETF during the sample period.

Robustness Check of Changes in Spreads

As a robustness check, the changes in spreads surrounding decimalization are reexamined by controlling for other variables that are known to affect spreads. It has been documented that bid–ask spreads might be affected by the securities' prices, volatility, and trading volumes (Stoll, 2000). In addition, the trading volume may also change due to decimalization.

Wang, Yau, and Baptiste (1997) find that volume and bid–ask spreads are jointly determined. First the Hausman (1978) test is conducted to investigate whether bid–ask spreads and volume are endogenous variables. The results of the Hausman test indicate that bid–ask spreads and volume are endogenously determined, and thus the two-stage least-squares method (2SLS) is employed to estimate the following regression model of the impact of decimalization on spreads:

$$m_{it} = \alpha_0 + \alpha_1 d_{it} + \beta_1 \log p_{it} + \beta_2 \log \sigma_{it} + \beta_3 \log V_{it} + \varepsilon_{it} \quad (12)$$

where m_{it} is the average spread measures of ETFs and of index futures at Day t, V_{it} is the average daily trading volume for ETFs and the average daily tick volume for index futures, respectively, σ_{it} is the daily volatility estimator, and p_{it} is the price at Day t divided by the average price.

The Parkinson (1980) extreme value estimator is applied as a proxy for the daily volatility, which is calculated as $\sigma_{it} = 0.361 \times [\log(H_t/L_t)]^2$, where H_t and L_t denote the high and low prices of Day t, respectively.[9] The instrumental variables for the trading-volume variable include the daily 3-month T-bill rate, the daily volatility, and the open interest for index futures at Day $t - 1$, which is applied to index futures only. The dummy variable d_{it} is equal to 1 for the postdecimalization period and is 0 for the predecimalization period. In general, a negative coefficient of α_1 in the bid–ask spread equation would indicate reductions in spreads after decimalization.

Information Transmission Test

Although the major approach for analyzing the lead–lag relation is based on the vector error-correction model (VECM), there are various other methods available in the literature. Fleming, Ostdiek, and Whaley (1996), for example, employ a multivariate regression approach with explanatory variables including the lead and lag variables and error-correction terms. There are many cases wherein high-frequency intraday data may contain missing observations because of unevenly spaced transaction data. DeJong and Nijman (1997) propose an estimator that takes into account the problem of missing observations. Their estimator has been applied in many studies, such as DeJong and Donders (1998); DeJong, Mahieu, and Schotman (1998); and Beaulieu et al. (2003). However, the irregularly spaced data problem does not seem a serious

[9]Alizadeh, Brandt, and Diebold (2002) show that such a range-based estimation of volatility is not only a highly efficient volatility proxy, but also robust to microstructure noises, such as the bid–ask bounces.

issue for the index instruments in this study, as the S&P 500, Nasdaq-100, and DJIA index instruments all display relatively superior liquidity. Hence, the analysis is based on the information-share approach that requires the estimation of the VECM model.

The process of price discovery is further analyzed by using the Hasbrouck (1995) model, which calculates information shares as relative contributions of the variance of a security in the variance of innovations of the unobservable efficient price. According to Hasbrouck, the efficient price v_t follows a random walk: $v_t = v_{t-1} + u_t$. The observed prices of several cointegrated markets contain the same random-walk component and components incorporating effects of market frictions.

The method relies on the estimation of the VECM:

$$\Delta p_t = \mu + \sum_{i=1}^{k} A_i \, \Delta p_{t-i} + \gamma \, z_{t-1} + \varepsilon_t \qquad (13)$$

where p_t is an $n \times 1$ vector of cointegrated prices, A_i are $n \times n$ matrices of autoregressive coefficients, k is the number of lags, $z_{t-1} = \alpha' p_{t-1}$ is an $(n-1) \times 1$ vector of error-correction terms, γ is an $n \times (n-1)$ matrix of adjustment coefficients, and ε_t is an $n \times 1$ vector of price innovations. The coefficients γ of the error-correction term measure the price reaction to the deviation from the long-term equilibrium relationship. In the present VECM, $\Delta p_t = (\Delta F_t \, \Delta S_t)'$, where F_t and S_t are the prices of corresponding index futures and ETFs, respectively.

Hasbrouck (1995) shows that the following vector moving average model (VMA) can be derived from the VECM:

$$\Delta p_t = \Psi(L)\varepsilon_t \qquad (14)$$

where $\Psi(L)$ is a polynomial in the lag operator. The VMA coefficients can be used to calculate the variance of the underlying efficient price:

$$\sigma_u^2 = \Psi \Omega \Psi' \qquad (15)$$

where Ψ is a raw vector composed of VMA coefficients and $\Omega = \mathrm{var}(\varepsilon_t)$.

With the use of the Cholesky factorization to transform Ω into a lower triangular matrix F, $\Omega = FF'$, the information share of market j is calculated as:

$$I_j = \frac{(\Psi F)_j^2}{\sigma_u^2} \qquad (16)$$

where (ΨF_j) is the jth element of the row matrix ΨF. A market's contribution to price discovery is measured as the market's relative contribution to the variance of the innovation in the common trend. Baillie, Booth, Tse, and Zabotina (2002) show an easier method of calculating information

shares directly from the VECM results without obtaining the VMA representation. The present calculation of the information share is based on their method.

The results of information shares typically depend on the ordering of variables in the Cholesky factorization of the innovation covariance matrix. The first (last) variable in the ordering tends to have a higher (lower) information share, and this discrepancy may be large if the series' innovations are highly and contemporaneously correlated. Thus, the order of ETFs and index futures are alternated in the calculation and treat the resulting estimates as the upper bound (lower bound) of an instrument's information share, when it is treated as the first (second) variable.

EMPIRICAL RESULTS

Changes in Spreads and Trading Activity

Table I shows the changes in spreads and trading activity for ETFs surrounding decimalization. The total numbers of the pre- and postdecimal tick-by-tick observations are 75,643 and 109,434 for S&P 500 ETFs, 182,796 and 202,233, for Nasdaq-100 ETFs, and 35,945 and 56,909 for DJIA ETFs, respectively. The *t* statistics are adjusted for heteroskedasticity and serial correlation with the use of the Newey and West (1987) procedure.

Spreads of ETFs generally decrease in the postdecimal period, which is similar to the findings for equity securities in the tick size change literature. (Bessembinder, 2003; Chakravarty, Wood, & Van Ness, 2004; Chung, Charoenwong, & Ding, 2004; Goldstein & Kavajecz, 2000; and others) For S&P 500 ETFs, the Roll's and CFTC spreads decrease significantly. For Nasdaq-100 ETFs, all spread measures except for the effective spreads decrease significantly. For DJIA ETFs, all spread measures drop significantly after decimalization.

For all ETFs, the quoted size at BBO decreases significantly after decimalization, which indicates that market depth at BBO decreases. However, except for the S&P 500 ETFs, the average volume per trade increases. This, combined with an overall increase in the average daily trading volume, shows that the trading activity of ETFs does not seem to be adversely affected by decimalization, albeit the market depth seems to be thinner. The variances of all ETF returns do not change significantly. No significant changes are found in the volatilities of ETFs after decimalization. Similarly, other studies have found inconsistent results of changes in volatility after decimalization. For example, Chakravarty et al. (2004)

TABLE I

Changes in Spreads and Trading Activity of ETFs

	Predecimal	Postdecimal	Difference
Panel A: S&P 500 ETFs			
Quoted spread (%)	0.1066	0.1084	0.0018
Effective spread (%)	0.0902	0.0881	−0.0021
Roll's spread (%)	0.0823	0.0781	−0.0042
CFTC spread (%)	0.0630	0.0534	−0.0096**
Volume per trade (shares)	4492.1243	4059.2229	−432.9013*
Average daily volume (million shares)	2.8893	3.6655	0.7762*
Quote size (100 shares)	5931.2431	3604.1456	−2327.1075**
Variance (×10⁻⁶)	3.9956	3.6183	−0.3773
Observations	75,643	109,434	
Panel B: Nasdaq 100 ETFs			
Quoted spread (%)	0.2155	0.1773	−0.0383*
Effective spread (%)	0.1643	0.1541	−0.0101
Roll's spread (%)	0.1086	0.0876	−0.0210**
CFTC spread (%)	0.0956	0.0763	−0.0193**
Volume per trade (shares)	9905.3743	11871.8943	1966.5200*
Average daily volume (million shares)	16.2394	20.2452	4.0058**
Quote size (100 shares)	134.7429	124.0707	−10.6722
Variance (×10⁻⁶)	26.8635	21.1293	−5.7362
Observations	182,796	202,233	
Panel C: DJIA ETFs			
Quoted spread (%)	0.1476	0.1408	−0.0068
Effective spread (%)	0.1208	0.1100	−0.0109*
Roll's spread (%)	0.1127	0.0966	−0.0161**
CFTC spread (%)	0.0883	0.0735	−0.0148**
Volume per trade (shares)	2351.7649	2675.8846	324.1197
Average daily volume (million shares)	0.7496	1.3540	0.6044*
Quote size (100 shares)	1286.7614	901.4723	−385.2892**
Variance (×10⁻⁶)	3.1742	3.3071	0.1329
Observations	35,945	56,909	

Note.   The predecimal sample period is from October 29, 2000 to January 28, 2001, and the postdecimal sample period is from January 29, 2001 to April 28, 2001. All variables are the tick-by-tick measures. Quoted spreads are calculated from the posted quotes and effective spreads are calculated from the signed trades and their corresponding quoted midpoint by the Lee and Ready algorithm. Roll (1984) spreads are calculated by a modified regression framework: $r_t = \alpha + \frac{1}{2}s \cdot \Delta Q_t + u_t$, where α is an intercept term and Q_t is a trade indicator. The CFTC spread measure, as suggested by Wang (1994), is calculated as the absolute value of the opposite changes in prices, which is employed by the CFTC. Volume per trade is the average number of shares per trade. Quote size is in round lots (100 shares) and is calculated as the sum of the quote sizes of the best bid and offer prices. The *t* statistics are adjusted for heteroskedasticity and serial correlation with the use of the Newey and West (1987) procedure.

*Significance level of 5%.
**Significance level of 1%.

find significant increases in volatility, but Bessembinder (2003) and Chung et al. (2004) document significant decreases in volatility.

From Table II, the total numbers of the pre- and postdecimal tick-by-tick observations for index futures are 1,714,422 and 2,294,999 for S&P 500 E-mini, 2,451,703 and 3,207,581 for Nasdaq-100 E-mini, and

Changes in Effective Spreads and Trading Activity of Index Futures

	Predecimal	Postdecimal	Difference
Panel A: S&P 500 E-mini			
Roll's spread (%)	0.0205	0.0217	0.0012
CFTC spread (%)	0.0200	0.0214	0.0014*
Daily tick changes	28105.2787	36428.5556	8323.2769**
Variance ($\times 10^{-8}$)	2.1990	2.4979	0.2989
Observations	1,714,422	2,294,999	
Panel B: Nasdaq 100 E-mini			
Roll's spread (%)	0.0338	0.0345	0.0007
CFTC spread (%)	0.0298	0.0324	0.0026*
Daily tick changes	40191.8525	50913.9482	10722.1317**
Variance ($\times 10^{-8}$)	9.9299	7.1804	−2.7495
Observations	2,451,703	3,207,581	
Panel C: DJIA Futures			
Roll's spread (%)	0.0442	0.0446	0.0004
CFTC spread (%)	0.0408	0.0413	0.0005
Daily tick changes	1142.3279	1440.6826	298.3547*
Variance ($\times 10^{-8}$)	23.5692	21.9794	−1.5898
Observations	69,682	90,763	

Note. The predecimal sample period is from October 29, 2000 to January 28, 2001, and the postdecimal sample period is from January 29, 2001 to April 28, 2001. Roll (1984) spreads are calculated by a modified regression framework: $r_t = \alpha + \frac{1}{2}s \cdot \Delta Q_t + u_t$, where α is an intercept term and Q_t is a trade indicator. The CFTC spread measure, as suggested by Wang (1994), is calculated as the absolute value of the opposite changes in prices, which is employed by the CFTC. The daily tick changes are the average trade price changes in a trading day. The *t* statistics are adjusted for heteroskedasticity and serial correlation with the use of the Newey and West (1987) procedure.
*Significance level of 5%.
**Significance level of 1%.

69,682 and 90,763 for DJIA futures, respectively. The Roll's spreads generally increase and the increase is significant for the S&P 500 E-mini futures. The CFTC spreads are significantly larger for S&P E-mini and Nasdaq E-mini futures. Thus, after decimalization, only the spreads of ETFs decrease, but those for the index futures do not. This may impact the price dynamics between ETFs and index futures. Similar to those of the ETFs, the variances of futures returns also do not change significantly surrounding decimalization.

It is interesting to note that the spreads of the index futures remain low relative to those of the ETFs, even after decimalization. As will be seen in later sections, this is likely one of the reasons why the index futures still assume a dominant role in the price-discovery process, although the ETFs gain some improvement in price efficiency. The phenomenon is consistent with the trading-cost hypothesis of Fleming et al. (1996).

Adverse Selection

The changes in trading costs due to decimalization may influence the instruments by which the informed traders trade. Four models of spread decompositions are extensively examined in order to determine directly whether informed traders trade ETFs more intensively after decimalization. The analyses in Table III are based on tick-by-tick observations and the sample sizes are the same as those in Table I. Panel A of Table III shows that the adverse selection component increases significantly for Nasdaq-100 ETFs and DJIA ETFs by the GH approach. From the GKN approach in Panel B, both the S&P 500 ETFs and the DJIA ETFs experience significant increases in adverse selection. From the HS approach in Panel C of Table III, the adverse-selection component of S&P 500 ETFs and the DJIA ETFs increase significantly. Finally, from Panel D,

TABLE III
Decomposition of ETF Spreads

	Adverse selection			Order processing		
	Predecimal	*Postdecimal*	*Difference*	*Predecimal*	*Postdecimal*	*Difference*
Panel A: GH model						
S&P 500 ETFs	0.2544	0.2504	−0.0040	0.7456	0.7496	0.0040
Nasdaq 100 ETFs	0.2950	0.3243	0.0293*	0.7050	0.6757	−0.0293*
DJIA ETFs	0.2098	0.2312	0.0214*	0.7902	0.7688	−0.2140*
Panel B: GKN model						
S&P 500 ETFs	0.3023	0.3283	0.0260*	0.6977	0.6717	−0.0260*
Nasdaq 100 ETFs	0.5350	0.5402	0.0052	0.4650	0.4598	−0.0052
DJIA ETFs	0.3207	0.3967	0.0760**	0.6792	0.6033	−0.0760**
Panel C: HS model						
S&P 500 ETFs	0.6156	0.6206	0.0050**	0.0299	0.0238	−0.0061**
Nasdaq 100 ETFs	0.1572	0.1466	−0.0106	0.0082	0.0040	−0.0042
DJIA ETFs	0.0748	0.0796	0.0048*	0.4635	0.4504	−0.0131**
Panel D: MRR model						
S&P 500 ETFs	0.2467	0.2743	0.0276**	0.6196	0.5728	−0.0468**
Nasdaq 100 ETFs	0.3941	0.4191	0.0250**	0.1942	0.1606	−0.0336**
DJIA ETFs	0.2012	0.2068	0.0056	0.6468	0.5516	−0.0952**

Note. The predecimal sample period is from October 29, 2000 to January 28, 2001, and the postdecimal sample period is from January 29, 2001 to April 28, 2001. The analyses are based on tick-by-tick observations and the sample sizes are the same as those in Table I. Four methodologies—Glosten and Harris (1988) (GH); George, Kaul, and Nimalendran (1991) (GKN); Huang and Stoll (1996) (HS); and Madhavan, Richardson, and Roomans (1997) (MRR)—are used to decompose the adverse selection component of ETF spreads. The *t* statistics are adjusted for heteroskedasticity and serial correlation with the use of the Newey and West (1987) procedure.

*Significance level of 5%.
**Significance level of 1%.

the MRR model also shows a general increasing pattern of the adverse-selection component of ETFs after decimalization. Consistent with expectations and with Gibson et al. (2003), the percentages of adverse-selection components of ETFs tend to increase after decimalization.

From the columns of the order-processing cost in Table III, it can be seen that the order-processing component generally decreases after decimalization. These results point to a general increase in the adverse-selection component and a decrease in the order-processing component of the ETFs' spreads. This provides evidence showing that informed traders are now trading ETFs more intensively, because of reduced minimum tick size and lowered trading costs.

Another interesting observation from Table III is that according to the GH, GKN, and MRR models, the Nasdaq-100 ETFs have the largest proportion of adverse selection component, and the S&P 500 and DJIA ETFs have smaller and roughly equal proportions of adverse selection components.[10] This is to be expected, because the Nasdaq-100 index is narrower than the S&P 500, and its component stocks are also mostly smaller in size than those of the DJIA and S&P 500 indexes.

Robustness Check of Changes in Spreads

As a robustness check, structural models are used to examine whether the postdecimalization changes in bid–ask spreads are robust. First the simultaneity of spreads and volume is examined. The spread measures employed here are the average daily CFTC spreads with a total of 124 observations.[11] Panel A of Table IV presents the Hausman test results. Except for the bid–ask spread equation of the DJIA ETFs, all test statistics are statistically significant at the 5% level, which indicates that volume and bid–ask spreads are simultaneously determined. Hence, a simultaneous equation model similar to that of Wang et al. (1997) is conducted to analyze the impact of decimalization on the bid–ask spread of ETFs and index futures.

Panel B of Table IV presents the estimation results of the 2SLS method for the volume and the spread equations. To be concise, only results of the bid–ask spread equation are reported, as the particular interest here is on the relative changes in the trading costs after decimalization.

---

[10]The adverse selection component of the HS model cannot be compared across different types of ETFs, as it is not calculated as a proportion of the spreads.
[11]For ETFs, the same analysis is also conducted with quoted and effective spreads. The results are qualitatively similar.

TABLE IV

Regression Results of the Hausman Test and Changes in Spread Measures

	S&P 500 E-mini	Nasdaq 100 E-mini	DJIA futures	S&P 500 ETFs	Nasdaq 100 ETFs	DJIA ETFs
Panel A: Hausman test results						
Volume equation	−237.700 (0.001)	−71.230 (0.012)	−37.210 (0.001)	6.991 (0.032)	6.818 (0.001)	−6.291 (0.043)
Bid–ask spread equation	−0.002 (0.031)	−0.006 (0.001)	−0.011 (0.002)	0.005 (0.047)	0.019 (0.001)	−0.003 (0.059)
Panel B: 2SLS estimation results of changes in bid–ask spread						
α_0	0.279 (0.001)	0.406 (0.043)	1.271 (0.001)	1.619 (0.043)	1.116 (0.081)	0.698 (0.060)
α_1	−0.002 (0.127)	−0.003 (0.161)	−0.001 (0.471)	−0.016 (0.012)	−0.022 (0.009)	−0.021 (0.056)
$\log p_{it}$	−0.026 (0.112)	−0.025 (0.001)	−0.107 (0.001)	−0.209 (0.050)	−0.053 (0.001)	−0.086 (0.062)
$\log \sigma_{it}$	0.0013 (0.001)	0.004 (0.002)	0.008 (0.001)	0.028 (0.045)	0.024 (0.003)	0.011 (0.101)
$\log V_{it}$	−0.003 (0.067)	−0.016 (0.010)	−0.024 (0.003)	−0.083 (0.087)	−0.051 (0.085)	−0.009 (0.175)
Observations	124	124	124	124	124	124

Note. This table reports the Hausman test and the 2SLS estimation results of the spread equation for the period of October 30, 2000 to April 28, 2001. The 2SLS equation is:

$$m_{it} = \alpha_0 + \alpha_1 d_{it} + \beta_1 \log p_{it} + \beta_2 \log \sigma_{it} + \beta_3 \log V_{it} + \varepsilon_{it}$$

where m_{it} is the CFTC bid–ask spreads of ETFs and index futures at Day t. V_{it} is the average daily trading volume for ETFs and tick volume for index futures, σ_{it} is the daily volatility estimator, p_{it} is the price at Day t divided by the average price, α_0 is the intercept, α_1 measures the effect of the dummy variable, and d_{it} is an indicator variable, which equals 1 for the post-decimalization period and 0 for the predecimalization period. Numbers in the parentheses are p values.

The three control variables are shown to have a significant impact on the bid–ask spreads of ETFs and index futures, and have commonly expected signs (Chung, Charoenwong, & Ding, 2004; Gibson et al., 2003; and others) The coefficients of the dummy variables representing the decimalization effect are all negative for ETFs. The estimated coefficients of S&P 500 ETFs and Nasdaq-100 ETFs are negative and significant at the 5% level. In contrast, the spreads of index futures do not change significantly after decimalization.

To summarize, the results in Table IV show reductions in trading costs for ETFs after other variables likely to affect the spread size are controlled for. It is important to note that there is no evidence showing changes in the bid–ask spreads of index futures, which confirms the results in Table I.

Information Transmission

According to the trading-cost hypothesis, securities with lower trading costs will lead in the price discovery process. From previous sections, it can be seen that after decimalization the trading costs of ETFs decrease significantly relative to those of index futures. This section tests whether ETFs assume a more important role in the price-discovery process.

Given the time-series nature of the data, an initial step in the information transmission analysis is to test whether each price series is stationary. For transmission tests, 5-minute interval data are retrieved from the tick data of ETFs and index futures. The total numbers of 5-minute observations for all ETFs and futures are 4,819 and 4,977 during the pre- and postdecimal periods, respectively. The traded prices of ETFs are generally scaled down by the exchange from their respective index levels, so that their prices per share are comparable to those of other stocks. To make the prices of ETFs comparable to those of index futures, the prices of ETFs are multiplied by an adjusting factor, and the adjusting factors for the S&P 500, Nasdaq-100, and DJIA ETFs are 10, 40, and 100, respectively.

The results of the Phillips-Perron (1988) unit root test for price levels indicate that the existence of unit roots cannot be rejected for all index instruments.[12] The VECM results show that for each group of the index instruments, the error-correction term is significant for the ETFs, but not for the index futures. Although this result does not seem to change after decimalization, the coefficients of the ETFs' error-correction terms decrease relative to those of the index futures. The changes in the short-run lead–lag relationship after decimalization are the most significant for the Nasdaq-100 and S&P 500 ETFs, which indicate that the relative strength of information transmission from ETFs to E-mini futures seems to increase after decimalization. The fact that the Nasdaq-100 and S&P 500 ETFs are the ones that exhibit the most significant improvement in the information leading role is also expected, because they both have much higher liquidity than the DJIA ETFs as shown previously in Table I.

Table V presents the information share of ETFs and index futures for the pre- and postdecimal periods. The column labeled "Upper bound" ("Lower bound") is the information share when the security is treated as the first (second) variable in the calculation, and the column labeled "Midpoint" reports the average of the upper bound and the lower bound of the information shares.

---

[12]To save space, the estimation results of unit-root tests and VECM are omitted, and they are available upon request.

TABLE V
Contributions to Price Discovery

	Information share			
	Upper bound	*Lower bound*	*Midpoint*	*Observations*
Panel A: S&P500 E-mini				
Predecimal	0.9845	0.8671	0.9258	4,819
Postdecimal	0.9706	0.9266	0.9486	4,977
Panel B: Nasdaq 100 E-mini				
Predecimal	0.9827	0.7849	0.8838	4,819
Postdecimal	0.9213	0.7030	0.8122	4,977
Panel C: DJIA Futures				
Predecimal	0.9909	0.6462	0.8186	4,819
Postdecimal	0.9493	0.5586	0.7539	4,977
Panel D: S&P 500 ETFs				
Predecimal	0.1329	0.0155	0.0742	4,819
Postdecimal	0.0734	0.0294	0.0514	4,977
Panel E: Nasdaq 100 ETFs				
Predecimal	0.2151	0.0173	0.1162	4,819
Postdecimal	0.2970	0.0787	0.1878	4,977
Panel F: DJIA ETFs				
Predecimal	0.3538	0.0091	0.1814	4,819
Postdecimal	0.4414	0.0507	0.2461	4,977

Note. The predecimal sample period is from October 29, 2000 to January 28, 2001, and the postdecimal sample period is from January 29, 2001 to April 28, 2001. Information shares of index futures and ETFs are calculated by the Hasbrouck (1995) model, which shows that the following vector moving average model (VMA) can be derived from the VECM: $\Delta p_t = \Psi(L)\varepsilon_t$, where $\Psi(L)$ is a polynomial in the lag operator. The VMA coefficients can be used to calculate the variance of the underlying efficient price: $\sigma_u^2 = \Psi\Omega\Psi'$, where Ψ is a raw vector composed of VMA coefficients and $\Omega = \text{var}(\varepsilon_t)$. With the use of the Cholesky factorization to transform Ω into a lower triangular matrix F, $\Omega = FF$, the information share of market j is calculated as: $I_j = (\Psi F)_j^2/\sigma_u^2$, where $(\Psi F)j$ is the jth element of the row matrix ΨF. A market's contribution to price discovery is measured as the market's relative contribution to the variance of the innovation in the common trend. The column labeled "Upper bound" ("Lower bound") refers to the information share when the instrument is treated as the first (second) variable in the model, and the column labeled "Midpoint" reports the average of the upper bound and the lower bound of the information shares.

From Table V, the midpoints of the information share of index futures are much higher than 50% in both subperiods, which indicates that index futures dominate in price discovery. Similar results have been found in the literature, and it is generally argued that futures contracts lead in price discovery, because of their lower trading costs and higher leverage effects. The E-mini index futures of S&P 500 and Nasdaq 100 have higher information shares than the regular index futures of DJIA, and this further shows that instruments with higher liquidity and lower trading costs are more likely to assume leading roles in price discovery. The results are consistent with Ates and Wang (2004), who show that the S&P 500 and Nasdaq-100 E-mini futures assume dominant roles in price discovery.

After decimalization, all the upper bounds of the index futures' information shares decrease and all the lower bounds of the ETFs' information shares increase. Except for S&P 500 ETFs, the midpoint and the upper-bound information shares of ETFs increase after decimalization. Consistent with other findings in previous sections, the efficiency of ETFs has increased, because after decimalization, there is a general increase in the relative contribution of ETFs to the price-discovery process. These results support the trading-cost hypothesis by Kawaller et al. (1987) and Fleming et al. (1996).

A dominant role of index futures in price discovery is still observed, despite their decreased information shares after decimalization. The dominant role of the index futures in price discovery is expected, as E-mini index futures still possess many trading advantages, such as a high leverage effect. Furthermore, ETFs market makers sometimes make bids and offers based on the price movements observed from the index futures.

SUMMARY AND CONCLUSION

The impact of reduction in trading costs due to decimalization on the price-discovery process between ETFs and index futures has been studied. ETFs started to trade in decimals on January 29, 2001, and index futures continued to trade in their original tick sizes. Because both ETFs and index futures are index instruments, the decrease in the minimum tick size of ETFs may have changed the price dynamics between ETFs and index futures. In order to provide additional evidence on the impact of changes in trading costs on the information transmission processes, index instruments that are informationally linked are studied. Furthermore, this study sheds lights on the influences of trading costs on how and where informed traders choose to exploit their information advantages.

Consistent with the decimalization literature of equity securities, it is found that for ETFs, the trading activity increases, but the market depth drops after decimalization. The spreads for ETFs generally become smaller, indicating that the implicit trading costs of ETFs decrease. Due to the trading-cost hypothesis, this is likely to induce informed traders to switch their trades from index futures to ETFs. This conjecture is confirmed by estimating four models of spread decompositions and show that the adverse selection components of ETFs generally increase significantly after decimalization. The adverse selection component is also found to be the largest for the Nasdaq-100 ETFs.

Based on these results, tests are conducted to determine whether the price-discovery process changes due to changes in spreads and their

components. It is found that ETFs start to lead index futures after decimalization, a phenomenon that is not observed before the event. Moreover, a general increase in the information share of ETFs is also found. Overall, ETFs experience decreases in trading costs and increases in information trading, and these in turn lead to an improvement in the role of ETFs in the price-discovery process. These results provide further support for the trading-cost hypothesis.

Finally, although decimalization improves the general informational efficiency of ETFs, its impact might be different for different types of market participants.[13] Further studies on how different types of market participants are affected by decimalization will reveal more information on the merit of decimalization. For example, a potential research topic related to the issue is how decimalization would impact the profitability of index futures arbitrages. This is a topic left for future research.

BIBLIOGRAPHY

Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. Journal of Finance, 57, 1047–1092.

Amihud, Y., & Mendelson, H. (1987). Trading mechanisms and stock returns: An empirical investigation. Journal of Finance, 42, 533–553.

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

Ates, A., & Wang, G. H. K. (2004). Information transmission in electronic versus open-outcry trading systems: An analysis of US equity index futures markets (working paper). Washington, DC: Commodity Futures Trading Commission.

Baillie, R. T., Booth, G. G., Tse, Y., & Zabotina, T. (2002). Price discovery and common factor models. Journal of Financial Markets, 5, 309–321.

Beaulieu, M. C., Ebrahim, S. K., & Morgan, I. G. (2003). Does tick size influence price discovery? Evidence from the Toronto Stock Exchange. Journal of Futures Markets, 23, 49–66.

Bessembinder, H. (2003). Trade execution costs and market quality after decimalization. Journal of Financial and Quantitative Analysis, 38, 747–777.

Bollen, N. P. B., Smith, T., & Whaley, R. E. (2003). Optimal contract design: For whom? Journal of Futures Markets, 23, 719–750.

Chakravarty, S., Wood, R. A., & Van Ness, R. A. (2004). Decimals and liquidity: A study of the NYSE. Journal of Financial Research, 27, 75–94.

Chung, K. H., Charoenwong, C., & Ding, D. K. (2004). Penny pricing and the components of spread and depth changes. Journal of Banking and Finance, 28, 2981–3007.

---

[13]Angel (1997) and Seppi (1997) argue that an extremely small tick size set by the regulatory agency may not necessarily be optimal for all types of traders.

De Jong, F., & Donders, M. W. M. (1998). Intraday lead-lag relationships between the futures, options and stock market. European Finance Review, 1, 337–359.

De Jong, F., Mahieu, R., & Schotman, P. (1998). Price discovery in the foreign exchange market: An empirical analysis of the Yen/Dmark rate. Journal of International Money and Finance, 17, 5–27.

De Jong, F., & Nijman, T. (1997). High frequency analysis of lead-lag relationships between financial markets. Journal of Empirical Finance, 4, 259–277.

Fleming, J., Ostdiek, B., & Whaley, R. E. (1996). Trading costs and the relative rates of price discovery in stock, futures, and option markets. Journal of Futures Markets, 16, 353–387.

George, T. J., Kaul, G., & Nimalendran, M. (1991). Estimation of the bid–ask spread and its components: A new approach. Review of Financial Studies, 4, 623–656.

Gibson, S., Singh, R., & Yerramilli, V. (2003). The effect of decimalization on the components of the bid–ask spread. Journal of Financial Intermediation, 12, 121–148.

Glosten, L. R., & Harris, L. E. (1988). Estimating the components of the bid/ask spread. Journal of Financial Economics, 21, 123–142.

Goldstein, M. A., & Kavajecz, K. A. (2000). Eighths, sixteenths, and market depth: Changes in tick size and liquidity provision on the NYSE. Journal of Financial Economics, 56, 125–149.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. Journal of Finance, 50, 1175–1199.

Hasbrouck, J., & Ho, T. S. Y. (1987). Order arrival, quote behavior, and the return-generating process. Journal of Finance, 42, 1035–1048.

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica, 46, 1251–1271.

Huang, R. D., & Stoll, H. R. (1996). Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. Journal of Financial Economics, 41, 313–357.

Kaul, G., & Nimalendran, M. (1990). Price reversals: Bid–ask errors or market overreaction? Journal of Financial Economics, 28, 67–93.

Kawaller, I. G., Koch, P. D., & Koch, T. W. (1987). The temporal price relationship between S&P 500 futures and the S&P 500 index. Journal of Finance 42, 1309–1329.

Lee, C. M. C., & Ready, M. J. (1991). Inferring trade direction from intraday data. Journal of Finance, 46, 733–746.

Madhavan, A., Richardson, M., & Roomans, M. (1997). Why do security prices change? A transaction-level analysis of NYSE stocks. Review of Financial Studies, 10, 1035–64.

Neal, R., & Wheatley, S. M. (1998). Adverse selection and bid–ask spreads: Evidence from closed-end funds. Journal of Financial Markets, 1, 121–149.

Newey, W. K., & West, K. D. (1987). A simple positive semi-definite, heteroscedasticity, and autocorrelation consistent covariance matrix. Econometrica, 55, 703–708.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. Journal of Business, 53, 61–65.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. Biometrika, 75, 335–346.

Roll, R. (1984). A simple implicit measure of the effective bid–ask spread in an efficient market. Journal of Finance, 39, 1127–1139.

Seppi, D. J. (1997). Liquidity provision with limit orders and a strategic specialist. Review of Financial Studies, 10, 103–150.

Stoll, H. R. (2000). Friction. Journal of Finance, 55, 1479–1514.

Van Ness, B. F., Van Ness, R. A., & Warr, R. S. (2001). How well do adverse selection components measure adverse selection? Financial Management, 30, 77–98.

Wang, G. H. K. (1994). An intraday analysis of bid–ask spreads and price volatility in the S&P 500 index futures market. Journal of Futures Markets, 14, 837–859.

Wang, G. H. K., Yau, J., & Baptiste, T. (1997). Trading volume and transaction costs in futures markets. Journal of Futures Markets, 17, 757–780.