Available online at www.sciencedirect.com

ScienceDirect

Journal of Financial Economics 82 (2006) 251–288

JOURNAL OF Financial ECONOMICS

www.elsevier.com/locate/jfec

Market transparency, liquidity externalities, and institutional trading costs in corporate bonds ☆

Hendrik Bessembinder[a,*], William Maxwell[b], Kumar Venkataraman[c]

[a]*David Eccles School of Business, University of Utah, Salt Lake City, UT, 84112, USA*
[b]*Eller School of Management, University of Arizona, Tucson, AZ, 85721, USA*
[c]*Cox School of Business, Southern Methodist University, Dallas, TX, 75275, USA*

Received 21 April 2005; received in revised form 8 September 2005; accepted 3 October 2005
Available online 18 May 2006

Abstract

We develop a simple model of the effect of public transaction reporting on trade execution costs and test it using a sample of institutional trades in corporate bonds, before and after initiation of the TRACE reporting system. Trade execution costs fell approximately 50% for bonds eligible for TRACE transaction reporting, and 20% for bonds not eligible for TRACE reporting, suggesting the presence of a "liquidity externality." The key results are robust to changes in variables, such as interest rate volatility and trading activity that might also affect execution costs. Market shares and the cost advantage to large dealers decreased post-TRACE.

☆The authors thank Yakov Amihud, Tihomir Asparouhova, Shmuel Baruch, Ekkehart Boehmer, Cliff Ball, Jeremy Graveline, Joel Hasbrouck, Mike Lemmon, Ananth Madhavan, Mark Shenkman, Hans Stoll, Rex Thompson, Christopher Vincent, Arthur Warga, an anonymous referee, and seminar participants at Vanderbilt University, the NBER Microstructure Conference, the University of Virginia, Claremont-McKenna College, Brigham Young University, Indiana University, Southern Methodist University, the University of Houston, the Stockholm School of Economics, the Northern Finance Association Meetings, and the University of Utah for valuable comments. Thanks are also due to Lehman Brothers and MarketAxess for the provision of data. The NAIC bond pricing data and Fixed Income Securities database are obtained from the Fixed Income Research Program of the Bauer College of Business at the University of Houston. The second author acknowledges financial support provided by Moody's Credit Market Research Fund. Earlier versions of this paper were titled "Optimal Market Transparency: Evidence from the Initiation of Trade Reporting in Corporate Bonds" and "Market Transparency and Institutional Trading Costs."

*Corresponding author.
E-mail address: finhb@business.utah.edu (H. Bessembinder).

doi:10.1016/j.jfineco.2005.10.002

These results indicate that market design can have first-order effects, even for sophisticated institutional customers.

JEL classification: D82; G14; G19

Keywords: Corporate bonds; Institutional trading costs; Market transparency

1. Introduction

Security markets vary greatly in their transparency, that is, in the amount of information regarding market conditions made public on a timely basis. Equity markets generally disseminate continuous pre-trade information, such as best quotations and, in some cases, information about unexecuted limit orders, and immediately report prices and sizes of completed trades. Most futures markets report trades, but do not disseminate pre-trade information. Foreign exchange markets disseminate only non-binding "indicative" quotations to the public, and do not report transactions at all. Corporate bond markets were traditionally similarly opaque, with quotations available only to a few market professionals, and with no public transaction reporting.

Market transparency has been the subject of a handful of studies, but neither the theoretical predictions nor the empirical evidence is conclusive as to whether market quality is enhanced by increased transparency. In this paper, we develop a simple model of how more precise estimation of values occasioned by increased transparency can affect trade execution costs, and test its implications by estimating institutional trading costs in corporate bonds when the National Association of Securities Dealers (NASD) began to publicly report transactions in approximately 500 corporate bond issues through its trade reporting and compliance engine (TRACE) on July 1, 2002. The initiation of TRACE transaction reporting provides a potentially powerful experiment for assessing whether transparency is important to market quality, because the corporate bond markets were quite opaque prior to TRACE. In contrast, previous empirical work on transparency studied relatively small changes in the already quite transparent equity markets.

Improved transparency has the potential to reduce trade execution costs for corporate bonds, for at least two reasons. First, transparency might reduce market-maker rents. Pagano and Roell (1996) emphasize that opaque markets tend to benefit relatively well informed dealers in their negotiations with customers. Also, transactions reported through the TRACE system can be monitored by self-regulatory agencies and by the United States Securities and Exchange Commission (SEC). The potential role of increased transparency in improving customers' ability to control and evaluate trade execution costs has been emphasized by Annette Nazareth, Director of the Division of Market Regulation of the SEC:

> For investors as well as regulators, the difficulty lies in establishing the prevailing market price for a bond. This generally is the base line that is used to assess whether a mark-up (trade execution cost) is reasonable… Improved transparency will enable investors to better determine the fair price of a bond. This will make them better able to protect themselves against unfair pricing…" (excerpted from testimony before the

United States Senate Committee of Banking, Housing, and Urban Affairs, June 17, 2004).

In addition to potential reductions in market-maker rents, improved transparency could decrease market-making costs. Naik et al. (1999) develop a model implying that improved transparency in a dealer market can improve inventory risk sharing, thereby decreasing inventory carrying costs.

Among the more important recent studies of bond markets, Green et al. (2004) and Harris and Piwowar (2005) examine trades in municipal bonds. Each study reports that small trades pay much larger percentage trading costs than large trades, and each set of authors conjectures that this may occur because unsophisticated small investors cannot readily evaluate the trading costs they pay in the opaque market for municipal bonds. However, neither study provides direct evidence on the relation between market transparency and trading costs.

This study provides direct evidence on the issue by analyzing trade execution costs for institutional (insurance company) transactions in corporate bonds before and after the introduction of transaction reporting for corporate bonds through TRACE. The results indicate average reductions in one-way trading costs for corporate bonds subject to TRACE transaction reporting of five to eight basis points. These estimated reductions in trading costs average 40–60% of pre-TRACE trading cost estimates, and equate to approximately $2,000 per trade in the present sample of insurance company transactions. Extrapolating beyond the present sample, we estimate market wide trading cost reductions of roughly $1 billion per year after the initiation of TRACE transaction reporting. The key empirical results are robust to the inclusion of control variables such as interest rate volatility and bond market trading activity that might alternately explain variation in trade execution costs. We also find that trading activity is less concentrated with large dealers and that the large dealer cost advantage previously documented by Schultz (2001) is reduced post-TRACE. Collectively, these results indicate that the public reporting of corporate bond trades has had first-order effects on market quality, even for the relatively sophisticated institutional traders that are the focus of this study.

In a contemporaneous paper, Edwards et al. (2006) examine determinants of cross-sectional variation in trade execution costs for corporate bonds using a comprehensive but proprietary database of transactions during 2003. Among other findings, they report that one-way transaction costs for those bonds whose trades are publicly disseminated through TRACE are one to four basis points lower, after controlling for other relevant factors. Similar point estimates are reported by Goldstein et al. (2005), who examine trading costs for BBB-rated bond issues during 2003, using a proprietary database provided by NASD.

A distinction between this analysis and that provided by Edwards et al., and Goldstein et al., is that we estimate the effect of TRACE reporting on bond market quality around the time that public reporting of trades through TRACE was *first initiated*, on July 1, 2002, while Edwards et al., and Goldstein et al., consider the impact of transparency on execution costs during 2003, after transaction reporting had already been introduced for about 500 bond issues. This distinction is important if transaction reporting for some bond issues also improves market quality for other issues. The model we develop implies that this should indeed be the case. Consistent with this reasoning, Amihud et al. (1997) document that an improvement in the trading mechanism used for a subset of Tel Aviv Stock Exchange securities led to enhanced liquidity not only for the affected stocks, but also for correlated stocks with no change in trading mechanism. They coined the phrase *liquidity externality*, which they attributed to "the fact that improved value discovery for one security facilitates value discovery for the other (correlated) security."

If, as our model implies, a similar liquidity externality exists for corporate bonds, then an analysis of trade execution costs after TRACE initiation will understate the importance of trade transparency in reducing transaction costs. A liquidity externality seems particularly plausible for corporate bonds, since market practitioners often estimate the value of non-traded bonds based on "matrix" pricing that incorporates bond characteristics and observed prices for bonds that do trade.[1] Improved information about market transactions in some bonds should allow more accurate valuation and better monitoring of trade execution costs for non-reported bonds as well.

Consistent with the existence of liquidity externalities, we document that one-way trading costs for *non*-TRACE-eligible bonds decreased by about three and a half basis points on average after transaction reporting through TRACE was initiated in July 2002. For non-TRACE-eligible bonds issued by firms in the same industry as a firm with at least one bond issue eligible for TRACE reporting, the reductions in one-way trading costs are larger, averaging about five basis points. More to the point, the estimated five to eight basis point reduction in trading costs for TRACE-eligible bonds reported here is substantially larger than the 2.1 basis point cross-sectional estimate for large trades reported by Edwards et al.

The difference in the point estimates between this study and Edwards et al., points out a broader issue in the literature, namely the use of matched firm or control samples in general. While a matched approach is usually the preferred methodology, in applications where the effects of a treatment spill over to the general population it will not be possible to measure the effect of the treatment by comparing treated and non-treated observations.

Finding larger trading cost reductions in the present sample is all the more striking since we measure trading costs for institutional transactions. If opaqueness is primarily a problem for naïve individual investors then we should observe little or no effect of TRACE reporting. In contrast, the substantial effects documented here support the conclusion that transparency is important to institutional customers as well.

This paper is organized as follows. Section 2 reviews recent papers on bond markets and on market transparency. In Section 3 we present a theoretical model that examines the potential role of transaction reporting on bond valuation and trading costs. Section 4 discusses methods for obtaining trading cost estimates for corporate bonds, while Section 5 describes the available data and some implementation issues that arise. Section 6 presents the key empirical results and tests of robustness, while Section 7 concludes.

2. The recent literature on bond markets and transparency

2.1. Recent studies of the bond markets

The increasing availability of data has spurred a substantial volume of recent research focused on bond markets. Hotchkiss and Ronen (2002) study a sample of 55 high-yield

[1] Both buy and sell side bond traders typically have matrix pricing information available on their computer displays. Sell-side traders can also view an automatic matrix valuation for any bond based on observed trades in bonds that are similar in terms of credit rating and maturity.

bonds. The source of their data was the fixed income pricing system (FIPS), a predecessor to TRACE, which disseminated hourly summary reports on the pricing of a select set of high-yield bonds. They focus on the relative information efficiency of stock and bond markets, reporting that stock price changes do not systematically lead bond price changes, and that market quality as measured by pricing errors in the Hasbrouck (1993) framework is similar across stocks and bonds.

Schultz (2001) provides estimates of trading costs for a large sample of corporate bonds. He obtains a dataset of insurance company trades in corporate bonds from Capital Access International (CAI). For the period January 1995 to April 1997 he estimates average round-trip trade execution costs of about 27 basis points. Schultz also reports that active institutions pay less than inactive institutions, which suggests that trading costs in the relatively opaque pre-TRACE bond markets depended in part on the customer's degree of sophistication and familiarity with the bond markets.

Chakravarty and Sarkar (2003) also study the CAI database of insurance company bond transactions and report average trading costs for municipal bonds of 23 basis points, compared to 21 basis points for corporate bonds and 8 basis points for Treasury bonds. They also report that in the cross-section, spreads rise with bond maturity and credit risk, and fall with trading volume. Hong and Warga (2000) compare insurance company trades in the CAI database to trades on the NYSE Automated Bond System, reporting similar bid-ask spreads on each. Chen et al. (2003) adapt the methodology of Lesmond et al. (1999) to examine the liquidity of corporate bonds, using proprietary data from Bloomberg, and document that their liquidity measure can explain 16% of the variation across bonds in yield spreads. This finding is potentially important because it implies that market liquidity not only determines transactions costs, but may also affect the valuation of the bonds themselves.

In a related study, Edwards et al. (2006) examine trading costs for corporate bonds. Their non-public sample includes all transactions reported to the NASD during calendar year 2003, including those made public through TRACE and those not made public. Edwards et al., document several important empirical regularities regarding corporate bond trading costs. First, they report that corporate bond trading costs also decrease with trade size, a result that the authors attribute to small traders' lack of sophistication in combination with limited transparency. Second, they provide evidence as to cross-sectional variation in corporate bond trading costs, documenting that costs increase with time from issue, and decrease with better credit rating, issue size, bond complexity, the presence of a floating interest rate and a previous private equity issue. Third, they address the role of transparency, reporting that in the cross-section trade execution costs are lower for bonds whose trades are publicly disseminated through TRACE, after controlling for variation in other characteristics. In particular, their Table 5 reports point estimates indicating that one-way trade execution costs are reduced by 0.9 basis points for \$10,000 trades, 2.9 basis points for \$20,000 trades, 3.8 basis points for \$100,000 trades, and 2.1 basis points for \$1 million trades. Edwards et al., also report broadly similar point estimates from a time-series experiment, with execution costs for a set of bonds phased into the public dissemination of trade reports during 2003 declining by about three to four basis points on average. Given the result obtained here that liquidity externalities exist for corporate bonds, the point estimates provided by Edwards et al., may be viewed as quantifying the effect of public dissemination of trade information through TRACE *conditional* on transaction prices for other bonds already being available through TRACE, and therefore providing a lower bound on the overall effect of TRACE reporting on corporate bond execution costs.

2.2. Studies of market transparency

Security market transparency refers to the amount of information regarding market conditions and transactions made public on a timely basis. Transparency is often categorized as pre-trade transparency, which concerns the dissemination of quotations or other indications of trading interest (such as unexecuted orders in the limit order book), and post-trade transparency, which concerns the dissemination of data about completed trades. Markets that disseminate little or no price data are referred to as being opaque, or non-transparent.

Madhavan (1995) models a competitive dealer market in which all dealers are required to publicly disclose trades, and compares the resulting equilibrium to that obtained when trade disclosure is not mandatory for some or all dealers. His analysis confirms the existence of an equilibrium where dealers choose to not disclose trades, because they profit from the associated reduction in price competition. These dealers will also post narrower spreads at the initiation of trading in order to attract informative order flow. Large informed and liquidity traders who spread their orders across multiple periods benefit from the lack of trade disclosure in this model, since they can transact with dealers who are unaware of their previous trades. However, since the model assumes a competitive market, dealers break even on average, and these gains come at the expense of small noise traders who arrive later.

Pagano and Roell (1996) focus on the effect of trade disclosure on the strategies of price setters (market makers or limit order traders). Their model implies that liquidity traders will pay lower average trading costs in the more transparent market, though not necessarily for all trade sizes. Naik et al. (1999) develop a model implying that improved transparency in a dealer market can improve inventory risk sharing, thereby decreasing inventory carrying costs. Consistent with these predictions, Flood et al. (1999) provide experimental evidence that pre-trade transparency reduces bid-ask spreads.

However, some theoretical analyses predict that less transparent markets might improve liquidity. In particular, Bloomfield and O’Hara (1999) argue that an opaque market might give market makers incentives to quote narrow bid-ask spreads, because the order flow attracted by narrow spreads contains valuable information about market fundamentals, and Bloomfield and O’Hara (2000) provide experimental evidence generally consistent with this reasoning. The empirical evidence from actual asset markets is inconclusive, in part because structural changes in the transparency of actual markets are rare. Gemmill (1996) examines the London Stock Exchange after two changes in required post-trade transparency for large block trades, and does not detect any change in liquidity. Madhavan et al. (2005) examine the liquidity of the Toronto Stock Exchange when during 1990 it began to publicly disseminate its limit order book, and document increased execution costs and greater price volatility after the increase in pre-trade transparency. Boehmer et al. (2005) examine the effect when the New York Stock Exchange began to disseminate limit order book information in January 2002. They document that limit order traders are able to use the information to refine their strategies and, in contrast to the findings of Madhavan et al., report improved liquidity as measured by transaction costs and the informational efficiency of prices. The inconclusive results obtained from these empirical studies could reflect that they studied changes in transparency in equity markets that were already quite transparent. The present analysis, in contrast, focuses on a market that was almost entirely opaque prior to the initiation of transaction reporting.

2.3. The transparency of the bond markets

The corporate bond market traditionally reported trades only to the parties involved, so investors could not compare their own execution price to other transactions. Even institutional investors had to invest significant time and effort to obtain market information, and were limited in their ability to compare their transaction prices to those of other investors. Limited information regarding current prices, in the form of "indicative" quotes, was available to institutional investors through a messaging system provided by Bloomberg. Investors could use this system to indicate interest in buying or selling a particular issue in an effort to solicit bids or offers, or could telephone dealers for quotes. The situation was even more difficult for individual investors, who were precluded from accessing virtually all real-time market information.

In an effort to bring greater transparency to the bond markets and provide additional regulatory oversight, the United States Securities and Exchange Commission (SEC) on January 31, 2001 approved rules requiring the NASD to report all over-the-counter secondary market transactions in a specified set of corporate bonds. The requirement initially applied to a set of 498 bonds with issuance size of \$1 billion or greater, and was implemented July 1, 2002.

During the 2002 sample period NASD members were required to report all corporate bond transactions to the TRACE system within 1 h and 15 min. For each trade the member is required to report bond identification (CUSIP or NASD symbol), the date and time of execution, trade size, trade price, yield, and a buy or sell indication. However, not all of the reported information is disseminated to the public: investors receive bond identification, the date and time of execution, and the price and yield for bonds specified as TRACE-eligible. Trade size is provided for investment-grade bonds if the par value transacted was \$5 million or less, otherwise an indicator variable denotes a trade of more than \$5 million. Investors can access the trade information on the NASD website without charge, but with a four-hour delay. The information is also retransmitted without delay via third-party vendors to subscribing investors. Institutional investors typically rely on a third-party vendor to disseminate the pricing information in an easily accessible and useable format, with MarketAxess apparently being the most widely used.

3. A simple model of market transparency and price efficiency

3.1. Estimating bond values

In this section, we present a model that examines how improved precision in valuation affects transaction costs in corporate bonds. The model considers the case of two bonds, but the intuition is readily generalized to the case of multiple bonds. Assume that the time t transaction price for bond j is

$$P_{jt} = \mathrm{E}_t(V_j) + \alpha S_j Q_{jt}, \tag{1}$$

where Q_{jt} is an indicator variable that equals 1 for buyer-initiated trades and -1 for seller-initiated trades, $\mathrm{E}_t(V_j)$ is the estimated value of bond j at time t, conditional on knowledge of Q_{jt}, and αS_j is the one-way transaction cost paid by customers to transact in bond j. More precisely and as discussed further in Section 4, αS_j is the non-information portion of the bid-ask spread, attributable to inventory costs, order processing costs and, possibly, market-maker rents. Bond value is estimated imperfectly, with valuation error defined as $\varepsilon_{jt} \equiv V_j - E_t(V_j)$, where $\varepsilon_{jt} \sim N(0, \sigma^2_{\varepsilon j})$. We posit that transaction costs increase with the variance of the valuation errors, for two complementary reasons. First, greater valuation errors increase the inventory-related risks of market making, for which risk-averse market makers must be compensated in equilibrium. Second, greater valuation errors can increase the market power of dealers, increasing the likelihood that economic rents can be extracted from less informed customers. The market maker in bond j observes noisy signals of value for bond j and also for bond k. Suppressing time subscripts, the signals are

$$Y_j \equiv V_j + U_j, \tag{2a}$$

$$Y_k \equiv V_k + U_k, \tag{2b}$$

where $U_j \sim N(0, \sigma^2_{Uj})$ and $U_k \sim N(0, \sigma^2_{Uk})$. We posit that improved market transparency due to TRACE reporting will improve the precision of these signals by reducing the variance of the noise components U_j and U_k. In particular, observing past transactions helps the dealer, as well as customers, make more accurate inferences about bond value. Let V denote the unconditional mean valuation for all bonds, and define individual bond valuation factors as

$$I_j \equiv V_j - V, \tag{3a}$$

$$I_k \equiv V_k - V, \tag{3b}$$

where $I_j \sim N(0, \sigma^2_{Ij})$ and $I_k \sim N(0, \sigma^2_{Ik})$. We assume that all random variables except I_j and I_k are independent of each other. The covariance between I_j and I_k, denoted $\sigma_{Ij,Ik}$, allows for the possibility that common factors affect the values of both bonds j and k.

Given joint normality, the projection theorem implies that a market maker's optimal estimate of value is obtained as a linear projection of value on observed signals. Letting β_{jj} and β_{jk} denote the weights that are placed on the bond j and bond k signals, respectively, the linear estimate of bond j value is

$$E(V_j) = \beta_{jj} Y_j + \beta_{jk} Y_k, \tag{4}$$

implying that the valuation error can be expressed as

$$\varepsilon_j \equiv V_j - \mathrm{E}(V_j) = \mathrm{V}(1 - \beta jj - \beta_{jk}) + I_j(1 - \beta_{jj}) - \beta_{jj} U_j - \beta_{jk}(I_k + U_k), \tag{5}$$

with variance:

$$\sigma^2_{\varepsilon j} = \sigma^2_{Ij}(1 - \beta_{jj})^2 + \sigma^2_{Uj}\beta^2_{jj} + (\sigma^2_{Ik} + \sigma^2_{Uk})\beta^2_{jk} - 2(1 - \beta_{jj})(\beta_{jk})\, \sigma_{Ij,Ik}. \tag{6}$$

Notably, the larger the covariance between bond values, the lower is the variance of the bond k price error, implying lower spreads when bond values have a greater common component, ceteris paribus. The projection theorem implies that the linear coefficients are given as

$$\beta_{jj} = A/B \tag{7a}$$

and

$$\beta_{jk} = 1 - A/B, \tag{7b}$$

where $A \equiv (\sigma^2_{Ij} + \sigma^2_{Ik} + \sigma^2_{Uk} - \sigma^2_{Ij,Ik})$ and $B \equiv A + \sigma^2_{Uj}$. Noting that B exceeds A by the noise in the bond j signal, σ^2_{Uj}, we observe that the bond j market maker optimally places increasing weight on the bond k value signal as the bond j signal becomes less precise. Further, an increase in the covariance of bond j and k values, $\sigma_{Ij,Ik}$, reduces the optimal weight in the own bond signal (β_{jj}) and increase weight in the related bond signal (β_{jk}).

Substituting (7a) and (7b) into (6), the variance of the valuation error can be expressed as

$$\sigma^2_{\varepsilon j} = A(1 - A/B) \tag{8}$$

3.2. *The effect of noise reductions due to TRACE*

As noted above, we posit that bond transaction reporting can affect bond j trading costs by improving the precision with which bond values are estimated, i.e., by reducing noise in the value signal for bond j, σ^2_{Uj}, and in the related bond k, σ^2_{Uk}. If transaction reporting through TRACE were to eliminate the noise in the bond j signal ($\sigma^2_{Uj} = 0$), then $A = B$ and $\sigma^2_{\varepsilon j} = 0$. In contrast, if TRACE eliminated the noise in the related bond k signal ($\sigma^2_{Uk} = 0$), then the terms A and B are each reduced, and $\sigma^2_{\varepsilon j}$ is reduced but remains positive. Knowing the exact value of the related bond k does not give the exact value of bond j, but it does reduce the variance of valuation errors.

However, we do not anticipate that TRACE reporting will entirely eliminate noise in valuation. Transaction prices are still made public with a delay, and in any case differ from underlying values due to bid-ask bounce. (The public TRACE data does not include a buy-sell indicator, so it is not possible to know with certainty whether a particular transaction price lies above or below bond value.) To examine the effect of a TRACE-induced *reduction* in noise, differentiate (8) with respect to σ^2_{Uj} and σ^2_{Uk}, respectively, to get

$$\partial\sigma^2_{\varepsilon j}/\partial\sigma^2_{Uj} = (A/B)^2 > 0, \tag{9a}$$

$$\partial\sigma^2_{\varepsilon j}/\partial\sigma^2_{Uj} = [1 - (A/B)]^2 > 0. \tag{9b}$$

The right sides of both (9a) and (9b) are positive, implying that trade reporting will reduce valuation errors and spreads for bond j, if either own bond j prices are reported (9a), or if related bond k prices are reported (9b). The latter is a formalization of the liquidity spillover argument.

Interestingly, the effect of bond j transaction reporting on bond j spreads is not necessarily stronger than the externality effect of bond k transaction reporting on bond j spreads. The latter effect can be greater if the ratio of A to B is sufficiently small. This ratio declines with the noise in the bond j signal, σ^2_{Uj}. In the case of a bond whose own value is estimated with a great deal of noise, it could actually be more beneficial to reduce noise in the signal on a related bond that is already valued more precisely than to reduce the noise in the own bond value signal.

Finally, the magnitude of the liquidity externality is greater if bonds are more similar in the sense that $\sigma_{Ij,Ik}$ is greater. To demonstrate this, differentiate the right side of (9b) with respect to $\sigma_{Ij,Ik}$ to obtain

$$\partial^2\sigma^2_{\varepsilon j}/\partial\sigma^2_{Uk}\partial\sigma_{Ij,Ik} = 4[1-(A/B)]^2/B>0, \tag{10}$$

which implies that the reduction in bond j valuation errors and spreads that come from a reduction in the noise with which bond k values are estimated is greater if the bond values are more closely related.

We summarize the implications of this model as follows. If (i) transaction costs in any individual bond are increasing in the variance of that bond's pricing errors, and (ii) TRACE reporting improves the precision of value estimates, then we have the following:

H1: Direct effect. TRACE transaction reporting will reduce transaction costs for bonds whose trades are disseminated through TRACE.
H2: Liquidity externality effect. TRACE transaction reporting for some bonds will reduce transactions costs for other bonds, and ceteris paribus the effect will be stronger for bonds with lower tracking error, i.e., higher covariance with TRACE-eligible bonds.

4. Measuring trading costs in bond markets

Most studies of trade execution costs have focused on equity markets, and are able to exploit the existence of reliable quotation databases to construct measures of quoted (ask price less bid price) and effective (trade price relative to quotation midpoint) spreads. In contrast, data on bid and ask quotations are not broadly available for bond markets. However, some bond transaction databases do indicate whether a dealer participated as a buyer or a seller. As a consequence, several studies, including this one, adopt variations of indicator variable regressions to estimate trade execution costs for bonds.

The indicator variable model used here is related to those suggested by Huang and Stoll (1997), Madhavan et al. (1997), and Schultz (2001). Let S denote the effective one-way spread, i.e., half the difference between the price at which dealers will sell a bond and the price at which they will purchase the bond, initially assumed to be constant. Consistent with prior theory, we assume that the spread contains an informational component, γS, and a non-informational component αS that can reflect inventory costs, order processing costs, and possible economic rents, where $\alpha \equiv (1-\gamma)$. As in Section 3, Q_t denotes an indicator variable that equals 1 if the time t trade is a customer buy and -1 if it is a customer sell, and P_t denotes a transaction price at time t, which equals the ask quote when $Q_t = 1$ and the bid quote when $Q_t = -1$. V is the unobservable true value of the bond, and $\mathrm{E}_t(V)$ is the market maker's estimate of bond value, conditional on knowledge of whether the time t trade is a customer buy or sell. Transaction prices are given by Eq. (1), reproduced here with the subscript identifying bond j suppressed:

$$P_t = \mathrm{E}_t(V) + \alpha S Q_t. \tag{1a}$$

The market maker's estimate of bond value is updated due to surprises in order flow, $Q_t^* \equiv Q_t - \mathrm{E}_{t-1}(Q_t)$, and due to public information revealed since the prior period, denoted η_t:

$$\mathrm{E}_t(V) = \mathrm{E}_{t-1}(V) + \gamma S Q_t^* + \eta_t. \tag{11}$$

Madhavan et al. (1997), Huang and Stoll (1997), and others have documented that order flow in equity markets is positively autocorrelated. To allow for this possibility we follow

Madhavan et al., and assume that order flow follows a simple AR1 process, so that $E_{t-1}(Q_t) = \rho(Q_{t-1})$. Letting $\Delta P \equiv P_t - P_{t-1}$ and $\Delta Q \equiv Q_t - Q_{t-1}$, the first difference of (1a) can be expressed as

$$\Delta P = \gamma S Q_t^* + \alpha S \Delta Q + \eta_t. \tag{12}$$

We assume that a fraction w of public information eventually becomes observable to econometricians in the form of data with realizations X_t, while the remaining portion is due to unobservable innovations U_t that represent statistical noise, so that $\eta_t = wX_t + (1-w)U_t$. Substituting into (12) gives

$$\Delta P = wX_t + \gamma S Q_t^* + \alpha S \Delta Q + (1-w)U_t \tag{13}$$

which is the basis for the regression specifications reported here.

In the special case where $\rho = 0$, so that order flow is not serially correlated, $E_{t-1}(Q_t) = 0$ and $Q_t* = Q_t$, allowing expression (12) to be restated as

$$\Delta P = \gamma S Q_{t-1} + S \Delta Q + \eta_t, \tag{14}$$

which is expression (5) in Huang and Stoll (1997). Our approach can therefore be viewed as a generalization of Huang and Stoll that allows for serial correlation in order flow and that incorporates observable public information.[2] Continuing to assume $\rho = 0$, expression (1a) can be combined with (11) to give

$$P_t = E_{t-1}(V) + SQ_t + \eta_t. \tag{15}$$

The bid quote, denoted B_t, is then given by (15) evaluated at $Q_t = -1$, and we have

$$P_t - B_t = (1+S)Q_t, \tag{16}$$

which is the basis for the estimation methods used by Schultz (2001) and Warga (1991).

Finally, relaxing the assumption that $\rho = 0$ so as to again allow for serial correlation in order flow, substituting $Q_t - \rho Q_{t-1}$ for Q_t* in (15) and rearranging terms gives

$$\Delta P = SQ_t - (S\rho\gamma + S\alpha)Q_{t-1} + \eta_t \tag{17}$$

which is equivalent to expression (4) in Madhavan et al. (1997).

4.1. Discussion

Expression (16) provides a simple method to estimate spreads, and has been implemented in bond markets by Schultz (2001) and Warga (1991). Aside from the maintained assumption that order flow is not serially correlated, the main drawback is that estimates of bid quotes are required. To implement this method, Schultz constructs estimates of bid quotes prevailing at the time of transactions. He obtains actual bid quotes as of the end of the prior month from the University of Houston Fixed Income database, and then adjusts the end-of-month quotes for changes in Treasury interest rates between the end of the prior month and the trade date.

[2]Note, however, that the coefficient obtained on ΔQ when implementing (14) estimates the entire half spread, while the coefficient obtained on ΔQ when estimating (13) is the non-informational component of the spread.

However, it might not be possible to implement this method in future studies. The need to adjust the monthly bid quotes to obtain within-month estimates caused Schultz to limit his study to investment-grade bonds. More importantly, the bid quote data disseminated by the University of Houston pertained only to the bonds contained in the Lehman Brothers Bond Indices, and has not been disseminated for transactions since 1999. The indicator variable model (13) in contrast can be implemented for all bonds, and does not require access to quotation data.

Relying on (13), we estimate regressions of the form

$$\Delta P = a + wX_t + \gamma SQ_t^* + \alpha S\Delta Q + \omega_t. \tag{18}$$

This specification generalizes the approaches of Huang and Stoll (1997) and Madhavan et al. (1997) to incorporate observable public information that affects bond value. Simulation-based evidence reported in the Appendix verifies that controlling for variation in factors that affect bond values improves the precision of estimates of the half-spread. This is particularly important for corporate bonds, since the elapsed time between trades can be long.[3]

We include in (18) three public information variables, each measured from the date of the most recent transaction on a day prior to the date of the current transaction. The first is the return on the on-the-run Treasury security matched to the corporate bond based on maturity.[4] Since the Treasury return control variable differs across bonds according to maturity, changes in the interest rate term structure are also accommodated. The second control variable is the change in the yield spread between BAA bonds and Treasury securities, included to allow for changes in market-wide risk perceptions and economic activity. The yield spread data are obtained from the Federal Reserve Statistical Release. The third control variable is the percentage return on the issuing firm's common stock, included to allow for company-specific news that could affect bond prices. Stock return data are obtained from the Center for the Research in Security Prices (CRSP) daily database. If the bond was issued by a subsidiary, we use the parent company for stock return data. Since relations between bond returns and both yield spread changes and common stock returns are likely to differ depending on the credit quality of the issuing firm, we use indicator variables to estimate distinct coefficients on these control variables for investment-grade (BBB-rated or better) and non-investment-grade bonds. We assess the impact of TRACE reporting by simply including the product of the ΔQ variable and an indicator variable (TRACE) that equals one for trades occurring after July 1, 2002 and zero for trades before.

[3]Note that we have implicitly assumed the bid and ask prices to be symmetric about estimated value, or equivalently, that dealer inventory does not affect quote placement. As the discussion in Huang and Stoll (1997) makes clear, if this assumption is relaxed the γ coefficient estimate can be interpreted to reflect the sum of trades' information content and a location parameter used to induced mean reversion in inventory. Our approach does allow inventory costs to affect spreads, as part of the non-information component α.

[4]On-the-run Treasuries are the mostly recent issued Treasury bonds in each maturity category. These are the most liquid treasury securities, and are typically used to define the Treasury yield curve. The identity of an on-the-run Treasury issue can change over time as the government issues new debt. In cases where the identify of the on-the-run security changes during the sample period, we keep the same benchmark Treasury issue as defined at the beginning of the year, but use the yield from the new on-the-run. Treasury security to determine the price of the matching Treasury issue.

5. Data description and implementation issues

5.1. Data sources and description

To examine trading costs before and after the implementation of TRACE, we rely on the National Association of Insurance Commissioners (NAIC) transaction data in corporate bonds. NAIC is also the source of the data used by Schultz (2001) and Campbell and Taksler (2003) who provide a more detailed description.

Schultz (2001) and Campbell and Taksler (2003) estimate that insurance companies hold between one-third and 40% of corporate bonds. Of course, insurance companies might not trade as actively as other market participants. A comparison of the NAIC data for TRACE-eligible bonds to the actual TRACE data (provided to us by MarketAxess) indicates that insurance companies completed 12.5% of the dollar trading volume in TRACE-eligible securities during the second half of 2002. So while the NAIC data are not exhaustive, they do represent a substantial portion of the corporate bond market. Importantly for purposes of this study, the NAIC data provide corporate bond transaction data both before and after the initiation of TRACE reporting.

NAIC data provide detailed transaction information, including trade date, price, size of the trade (market and face value), issue CUSIP, dealer identification, and the type of selling institution. Notably, the NAIC data do not contain transaction times. The effect of this omission is discussed further in Section 4.2. To obtain information on the characteristics of each traded bond, including maturity date and bond rating, we use the fixed income security database (FISD).

We divide the NAIC data into two samples: TRACE and non-TRACE. The TRACE sample consists of bond issues for which the NASD began to report transaction information on July 1, 2002, while the non-TRACE sample consists of the remaining issues in the NAIC sample. In selecting the time interval to study, we want to span enough time to provide accurate measures of trading costs, but also minimize the possibility of other factors influencing results. In addition, the post sample should include a long enough time frame for participants to become accustomed to the TRACE system. For the main analysis, the pre-TRACE period is defined as the six months prior to the implementation of TRACE, January 1, 2002 through June 30, 2002, and the post-TRACE period is the following six months, July 1, 2002 through December 31, 2002.

The TRACE sample begins with 51,209 transactions, representing all NAIC reported transactions in 2002 for the sample bonds eligible for TRACE reporting in July 2002. We eliminate transactions if no matching CUSIP could be obtained from CRSP (e.g., for foreign issuers or privately owned equity), which reduced the sample to 48,627 observations and a final sample of 439 TRACE-eligible bonds. We also eliminate sell transactions that involved the bond issuer, including those with any of the following terms in the transaction name field: called, cancelled, conversion, direct, exchanged, issuer, matured, put, redeemed, sinking fund, tax-free exchange, or tendered. We exclude transactions in which the dealer descriptions indicate the trade to be with a related party, as well as transactions labeled "no broker" and "private." These screens eliminate 2,166 transactions, leaving 46,461 observations.

The NAIC data could also suffer from data entry errors, as reports are manually coded. Prior researchers have handled this in different ways. Schultz (2001) discards observations that differ by more than 5% from the beginning and end-of-month bid price. Campbell and Taksler (2003) eliminate the top and bottom 1% of spreads from their analysis. Krishnan et al. (2004) eliminate all "inconsistent or suspicious" observations. We eliminate "reversal" transactions, where a given price exceeds *both* the preceding and following prices by at least 15%, or is less than both prices by the same magnitude. We also drop 5,533 trades that were either the first trade for the bond in the pre- or post-TRACE periods, or where matching stock and Treasury returns cannot be obtained for the TRACE reported date (usually a weekend or a holiday). We eliminate 1,341 trades where the absolute bond return exceeds 10%. The indicator regression specification (18) relies on the ability to identify "bid-ask bounce" in the series of transaction prices. Price changes exceeding 10% are not plausibly attributable to bid-ask spreads, but increase statistical noise. The final TRACE-eligible sample comprises 39,040 observations.

Identical screens are used for the non-TRACE sample. We begin with 94,400 NAIC transactions during 2002 for bonds that are not TRACE-eligible as of July 1, 2002. The requirement to match these bonds with both the FISD (bond characteristic) and CRSP databases reduces the sample to 82,647 trades. Eliminating government, quasi-government, and municipal debt securities and implementing the same filters as for TRACE bonds reduces the sample to 54,601 observations. We also delete bonds that are not in the sample for all of the year 2002, leaving a final non-TRACE sample of 53,282 transactions.

To examine industry-specific effects we obtain three-digit SIC codes for the issuer of each sample bond from CRSP. We then identify those non-TRACE bonds with issuers in the same three-digit SIC code as the TRACE bonds. Since TRACE eligibility is bond and not firm specific, we also create an indicator variable that equals one if an issuer matched on the basis of the six-digit CUSIP also has at least one TRACE-eligible bond.

5.2. Estimating trade execution costs without transaction times

As noted above, the NAIC database contains transaction dates, but not transaction times. The econometric model (18) relies on the assumption that transactions are appropriately ordered in time. The available information does not allow us to verify whether this is the case for transactions in the NAIC database.

We adopt the following procedure, which exploits the fact that the dataset contains transaction dates, if not transaction times. For each trade, we compute the change as the current observation minus the last observation contained in the dataset (which might or might not be chronologically last) on the most recent prior trading day. Eq. (18) is then estimated using the resulting data.

To our knowledge, the effect of imperfect time ordering of data observations on the properties of estimates obtained from models similar to (18) has not been the subject of analytical study. We therefore conduct a simulation analysis in which the underlying parameters are known, and assess whether the approach we use, as well as some plausible alternatives, provides unbiased estimates. Results of the simulation, reported in the Appendix, support the conclusion that, although the lack of time stamps reduces statistical power, the method used here provides unbiased parameter estimates.

However, this method also creates overlapping dependent variables on days when a bond trades more than once, which complicates statistical inference. Further, the number of adjacent observations that overlap is a random variable, equal to the number of trades on each day. To our knowledge no established procedure exists to compute consistent standard errors in a datasets with a random number of overlapping observations.[5] We therefore compute probability values for each coefficient estimate using a technique known as the ''block bootstrap'' that, unlike the standard bootstrap approach (which assumes independence across observations), relies on no specific assumption regarding the structure of the data-generating process. For descriptions of the block bootstrap approach see, for example, Carlstein (1986) and Hall and Jing (1996).

Like the standard bootstrap approach (see, e.g., Efron and Tibshirani, 1993), the block bootstrap relies on drawing observations from the original sample with replacement. The statistical model is estimated once for each bootstrap sample. By repeating the process a large number of times we create a distribution of bootstrap coefficient estimates. However, instead of drawing single observations, we draw blocks of consecutive observations to capture the dependence structure of neighboring observations.

We implement the block bootstrap using bond-days to define blocks. To ensure robustness we also estimate results using a week of trading for each bond to define a block; P-values are almost indistinguishable from those reported. Thus, if the original sample contains N bond-days with trades, each bootstrap sample is created by drawing N bond-days at random and with replacement from the original sample. We create 1,000 bootstrap samples and estimate (18) in each, leading to a distribution of 1,000 bootstrap sets of coefficient estimates. Note that this procedure not only accommodates the dependence of the observations within a trading day, but since a given trading day will appear more or less frequently across the bootstrap samples, it also accommodates any commonality in bond price movements attributable to different bonds trading on the same day. To assess the bootstrap probability value for a coefficient estimated from the actual data, we examine the proportion of the bootstrap estimates that are of the opposite sign as the actual estimate. For example, if the actual estimate is positive, but 121 of 1,000 bootstrap estimates are negative, the bootstrap p-value for coefficient is 0.121.

For our basic analysis we estimate (18) in two stages. We first estimate the autocorrelation in order flow by an OLS regression of the trade indicator variable Q_t on the trade indicator variable for the last trade recorded in the same bond on a prior day,

$$Q_t = c + \rho Q_{t-1} + \ddot{\upsilon}, \tag{19}$$

and use the resulting estimate of ρ to construct the variable $Q_t{}^*$, the surprise in order flow. Since the elapsed time between trades can vary greatly we estimate Eq. (18) by weighted least squares, where the weights are the inverse of the elapsed time (in days, plus one) between trades. Probability values are assessed by the block bootstrap method.

The procedure outlined here relies on weighted least squares to allow for heteroskedasticity and the block bootstrap to allow for overlapping observations, and seems likely to provide a reliable basis for inference. To ensure robustness, we also report results of some key specifications that are obtained when the parameters of expressions (18) and (19) and their standard errors are estimated simultaneously using the generalized method of moments (GMM) approach. As Madhavan et al. (1997) note, the GMM method is useful for estimating indicator variable models, since it allows both for generalized heteroskedasticity and, by use of the Newey–West covariance matrix estimate, for

[5]However, ours is not the first study in which this difficulty has arisen. Several studies of trading costs in equity markets, including Madhavan and Cheng (1997) and Bessembinder and Venkataraman (2004), estimate costs by comparing prices for individual trades to a common prior benchmark, such as the prior-day close. This also gives rise to a random number of overlapping observations. To our knowledge this study is the first to make an explicit attempt to address the resulting difficulties in inference.

autocorrelation in regression errors. The main shortcoming of the GMM method in this application is that the number of lags specified for the autocorrelation correction must be a constant, while the number of overlapping observations in our data is random. We assess the effect of altering the fixed lag length and find that our conclusions are robust. We implement the GMM method while using two (the median number of overlapping observations on days when a bond trades more than once) lags to allow for autocorrelation in the regression errors.

6. Empirical results

6.1. Descriptive data

Table 1 reports summary statistics for the TRACE, non-TRACE, and combined TRACE and non-TRACE samples. Panel A provides information about the characteristics of the bonds in the samples. It is evident that the TRACE-eligible bonds are larger ($1.45 versus $0.34 billion average issue size) and of higher credit quality than non-TRACE bonds. Information about trading characteristics is provided in Panel B. Sample bond prices were close to par during 2002, but were on average higher after TRACE.

The corporate bonds in the sample trade relatively infrequently. For the TRACE sample, the average number of insurance company trades is 46 per issue in the six months prior to TRACE and 50 per issue in the six months after TRACE implementation. Non-TRACE bonds trade less, with an average of 11 and 13 trades per issue pre- and post-TRACE implementation. The trades are relatively large, averaging $3.0 and $2.5 million pre-TRACE for TRACE- and non-TRACE-eligible bonds respectively. Average trade size increased post-TRACE, to $3.1 million for the TRACE sample and $2.9 million for the non-TRACE sample. The increased trade size post-TRACE indicates that orders are not split into smaller trades post-TRACE, as might be expected if liquidity supply had become scarce. Median trade sizes are $1 million or less, indicating positive skewness in trade sizes.

The NAIC database covers some $263 billion in bond trading during 2002, including $119 billion in TRACE-eligible issues and $144 billion in non-TRACE issues. In the second half of 2002, volume in the sample of TRACE-eligible bonds was $64.5 billion, and for comparison, the total volume (all transactions reported to the NASD under TRACE rules) for these bonds was $514 billion over the same time period. Consistent with the institutional nature of the NAIC data, the mean (median) transaction size in the NAIC sample is $3.1 million ($785 thousand), as compared to $508 thousand ($26 thousand) for all TRACE trades.

6.2. The effect of TRACE reporting on TRACE-eligible bonds

Table 2 reports the results of estimating specifications (18) and (19) for the sample of insurance company trades in TRACE-eligible bonds. The dependent variable in (18) is the price change in percent, so coefficient estimates can be interpreted in basis points.

The coefficient estimate on Q_{t-1} in (19) is positive and significant in all specifications. This implies that, consistent with the prior evidence from equity markets, order arrival in bonds is positively serially correlated. The positive autocorrelation in equity market order flow likely derives in part from the splitting of large orders into smaller orders executed

Table 1
Bond descriptive information
This table provides descriptive information for the three samples used in the paper: TRACE-eligible bonds, non-TRACE bonds, and TRACE and non-TRACE bonds. Panel A provides information about the bonds in the sample and Panel B provides information about trading volume characteristics.

	TRACE Bonds		Non-TRACE Bonds		TRACE & Non-TRACE Bonds	
Panel A: Descriptive bond information						
Total number of bond issues	439		3,122		3,561	
Average time to maturity (in years)	8.16		10.22		9.986	
Average issue size (in $M)	1,447		336		462	
Issue size						
Large (greater than $500 M)	423		578		1002	
Small (less than $500 M)	16		2,544		2,560	
Credit quality						
Investment-grade (BBB- thru AAA)	389		1,981		2,370	
Non-investment-grade (below BBB-)	50		1,141		1,191	
	Pre-TRACE	Post-TRACE	Pre-TRACE	Post-TRACE	Pre-TRACE	Post-TRACE
Panel B: Transaction price, transaction volume and transaction frequency						
Average trade price (% of par value)	99.90	101.87	98.66	102.29	99.19	102.11
Average number of trades by issue	46	50	11	13	17	19
Trade size						
Average trade size (in $MM)	2.98	3.09	2.48	2.93	2.69	2.99
Median trade size (in $MM)	0.81	0.78	0.88	1.00	0.85	0.96
Total number of trades	18,180	20,860	24,528	28,754	42,708	49,614
Cumulative trading volume (in $MM)	54,091	64,522	60,812	83,984	114,885	148,346

successively. Finding positive autocorrelation in order flow here is suggestive that large bond orders could also be split. The coefficient estimate on the surprise in order flow Q_t* in (18) is not significant. Since this coefficient estimates the effect of order flow on estimated bond value, this result is consistent with notion that the insurance companies in the NAIC sample trade for liquidity reasons, and not on the basis of private information.

The estimated coefficients on the control variables that measure public information flow in (18) are highly significant, suggesting that the inclusion of these variables improves the precision of the trading cost estimates. Coefficient estimates on stock returns are positive (bootstrap p-value $= 0.000$), consistent with the reasoning as well as with empirical evidence reported by Hotchkiss and Ronen (2002), that both stock and bond returns respond to new information about the value of the issuing firm's underlying assets. As might be expected, the coefficient on stock returns is greater when explaining returns on non-investment-grade bonds. The return on the benchmark Treasury bond also enters with a positive coefficient estimate, as both Treasury and corporate bonds respond to market-wide interest rate movements. The coefficient estimates on the change in the BAA-Treasury yield spread are negative and significant for non-investment-grade bond issues, but are generally close to zero for investment-grade bond issues. The broad significance of

Table 2
Spreads on TRACE Bonds
Panel A: Full sample half-spread estimates
In this table, we examine the half-spread for TRACE-eligible corporate bonds during 2002. We estimate a two-stage model with the first stage estimated as

$$Q_t = a + bQ_{t-1} + \varepsilon_t. \quad (1)$$

Referring to ε_t from Eq. (1) as Q_t*, the second stage is then estimated as,

$$\Delta P = a + wX_t + \gamma SQ_t^* + \alpha S\Delta Q + \omega_t. \quad (2)$$

In the regression, we include three public information variables, each measured from the date of the most recent transaction on a day prior to the date of the current transaction. The first is the change in the interest rate for an on-the-run Treasury security matched to the corporate bond based on maturity. The second is the percentage return on the issuing firm's common stock. The third is change in the spread between long-term indexes of BAA-rated bonds and U.S. Treasury securities. We then interact these factors with investment- and non-investment-grade indicator variables to account for potential differences in sensitivity based on the bond's risk. In Columns 1 and 3, we estimate the half-spread for TRACE-eligible bonds in 2002. To assess the impact of TRACE reporting, we interact the ΔQ_{it} variable with a TRACE indicator variable that equals one for trades occurring after June 30, 2002 and zero for earlier trades in Columns 2 and 4. In Columns 1 and 2, we estimate Eq. (2) using weighted least squares (WLS), with the weight being a function of the time between Q and Q_{t-1} and the statistical significance determined from bootstrapped probability estimates. In Columns 3 and 4, we estimate Eqs. (1) and (2) jointly using generalized method of moments (GMM) while controlling for serial correlation (at two lags) and heteroskedasticity.

Estimation technique Column #	WLS (1)	WLS (2)	GMM (3)	GMM (4)
First-stage results				
Intercept	0.0884***	0.0884***	0.2031***	0.2035***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Q_{t-1}	0.0785***	0.0785***	0.1800***	0.1799***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Second-stage results				
Intercept	−0.0079	−0.0080	0.0182	0.0182
(probability)	(0.213)	(0.206)	(0.170)	(0.169)
Q*	−0.0129	−0.0129	0.0033	0.0033
(probability)	(0.387)	(0.387)	(0.876)	(0.877)
Treasury Return	0.2220***	0.2210***	0.2496***	0.2493***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Stock return × investment	0.0498***	0.0500***	0.0467***	0.0472***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Stock return × noninvestment	0.0801***	0.0798***	0.0822***	0.0820***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
ΔBaaTrsy × investment	−0.0005	−0.0003	−0.0763	−0.0005
(probability)	(0.210)	(0.229)	(0.804)	(0.858)
ΔBaaTrsy × noninvestment	−0.0447***	−0.0450***	−0.0448***	−0.0448***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Delta$Q$	0.0926***	0.1352***	0.0916***	0.1182***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Delta$Q$ × TRACE		−0.0789***		−0.0491**
(probability)		(0.002)		(0.016)
Adjusted R^2	3.18%	3.23%	3.56%	3.61%
N	38,970	38,970	38,970	38,970

***, **, *Denote statistical significance at the 99%, 95%, and 90% level, respectively.

Table 2 (*continued*)

Panel B: Half-spreads by volume and credit rating

In this panel, we examine the impact of TRACE on TRACE-eligible bonds' half-spread based on trading volume in Column 1 (bonds are segmented into terciles based on volume), and credit rating in Column 2 and 3 (if the firm is investment [BBB- and above] or non-investment [below BBB-] grade). The basic model is described in more detail in the Table 2, Panel A heading. We estimate the second stage using weighted least squares (WLS), with the weight being a function of the time between Q and Q_{t-1}. Reported in parentheses are bootstrapped probability values.

	By volume	By credit rating	
Column #	High vol./low vol. (1)	Investment-grade (2)	Non-investment-grade (3)
First-stage results			
Intercept	0.0884***	0.0987***	−0.0819***
(probability)	(0.000)	(0.000)	(0.000)
Q_{t-1}	0.0785***	0.0760***	0.0919***
(probability)	(0.000)	(0.000)	(0.000)
Second-stage results			
Intercept	−0.0082	−0.0229	0.2500**
(probability)	(0.212)	(0.486)	(0.028)
Q^*	−0.0128	−0.0109	0.0679
(probability)	(0.389)	(0.378)	(0.120)
Treasury return	0.2217***	0.2159***	0.2842***
(probability)	(0.000)	(0.000)	(0.000)
Stock return × investment	0.0500***	0.0497***	
(probability)	(0.000)	(0.000)	
Stock Return × noninvestment	0.0799***		0.0809***
(probability)	(0.000)		(0.000)
ΔBaaTrsy × investment	−0.0007	0.0002	
(probability)	(0.204)	(0.284)	
ΔBaaTrsy × noninvestment	−0.0443***		−0.0555***
(probability)	(0.000)		(0.000)
Delta$Q$		0.1292***	0.1565**
(probability)		(0.000)	(0.015)
Delta$Q$ × TRACE		−0.0635**	−0.3603***
(probability)		(0.028)	(0.001)
Low volume × Delta$Q$	0.3313***		
(probability)	(0.000)		
Low Vol. × Delta$Q$ × TRACE	−0.0730*		
(probability)	(0.079)		
Mid volume × Delta$Q$	0.1308***		
(probability)	(0.000)		
Mid volume × Delta$Q$ × TRACE	−0.0878**		
(probability)	(0.021)		
High volume × Delta$Q$	0.1151***		
(probability)	(0.000)		
High vol. × Delta$Q$ × TRACE	−0.0673**		
(probability)	(0.047)		
Adjusted R^2	3.33%	2.48%	10.28%
N	38,970	36,806	2,163

***, **, *Denote statistical significance at the 99%, 95%, and 90% level, respectively.

the control variables underscores the usefulness of controlling for changes in public information that affect corporate bond prices.

Over the full time period the coefficient on ΔQ, which estimates one-way trade execution costs for the institutional bond trades, is 9.3 basis points using the WLS approach (Column 1) and 9.2 basis points using the GMM approach (Column 3). As noted in Section 4, the coefficient on ΔQ more precisely estimates the non-information component of the half-spread. However, given that estimates of the information component obtained in this study do not differ significantly from zero, the distinction is immaterial when discussing our results. Columns 2 and 4 in Panel A reports results of a specification that includes the product of an indicator variable that equals one for trades after July 1, 2002 and zero otherwise and the change in trade indicator variable in order to estimate the impact of market transparency. The pre-TRACE estimate of the half-spread using the WLS (GMM) approach is 13.7 (11.8) basis points, which corresponds quite closely to the Schultz (2001) estimate of 27 basis points for the full spread in his study of insurance company trades in high credit quality corporate bonds over the 1995–1997 period. The similarity of the estimates reported here for the first half of 2002 and the estimates reported by Schultz for 1995–1997 suggests that corporate bond trading costs were relatively stable prior to the introduction of TRACE. Moreover, the similarity in point estimates obtained while using markedly different estimation procedures increases confidence in the reliability of the specification.

The point estimates reported in Panel A of Table 2 indicate substantial reductions in estimated trade execution costs for the sample of institutional corporate bond trades after TRACE reporting was initiated. The estimated trading cost reduction obtained from the WLS specification is 7.9 basis points, which is 58% of the corresponding pre-TRACE estimate. The GMM specification indicates a trading cost reduction of 4.9 basis points, which is 42% of the associated pre-TRACE estimate. A five-to-eight basis point decrease in trade execution costs equates to transaction cost savings ranging from \$32 to \$51 million (\$64.5 billion in trading multiplied by 0.00049 or 0.00079) during the last half of 2002 for the insurance companies in the present sample alone.

The estimated decrease in trade execution costs after TRACE initiation reported here is substantially larger than the estimate of 2.1 basis points for million dollar trades (the largest trade size reported) obtained in the cross-sectional analysis of 2003 bond trading presented by Edwards et al., However, the estimates of the levels of trading costs post-TRACE are similar, though somewhat smaller (as might be expected in this sample of exclusively institutional trades) than those reported by Edwards et al.[6] We conjecture that our estimates of trading cost reductions due to TRACE are greater, even while estimates of levels of trading costs are similar, because a liquidity externality improved market quality for all corporate bonds, including those not reported through TRACE. Before investigating this conjecture, we provide some evidence on cross-sectional variation in trading cost reductions for TRACE-eligible bonds.

Panel B of Table 2, Column 1 reports results obtained when separate trading cost estimates are obtained for liquid and illiquid bond issues. To assess liquidity we simply

[6]For example, their sample C1 in Table 6 is broadly similar to our TRACE-eligible sample. We estimate post-TRACE trading costs of 5.6 (WLS) or 6.9 (GMM) basis points, compared to their estimate of 8.8 basis points for large trades. Their samples T and C3 in Table 6 are broadly similar to our non-TRACE sample. They estimate trading costs of 20.1 and 20.8 basis points, compared to our estimates (Table 4, Panel A) of 16.5 (WLS) or 18.6 (GMM) basis points for the non-TRACE sample after TRACE reporting was initiated.

count the number of trades during the pre-TRACE sample period, and assign bonds to volume terciles based on transaction frequency. The results indicate that illiquid bonds paid considerably larger trade execution costs pre-TRACE (33.1 basis points for low volume, 13.1 basis points for mid volume, and 11.5 basis points for high volume). All three terciles saw roughly similar decreases in execution costs post-TRACE (7.3 basis points for low volume, 8.8 basis points for mid volume, and 6.7 basis points for high volume).

Columns 2 and 3 of Panel B provide estimates of trade execution costs for TRACE-eligible bonds as a function of bond credit rating. We report the results as separate subsamples given the large difference across subsamples in the coefficient on the Treasury return variable. The results indicate larger execution costs pre-TRACE for non-investment-grade bonds (defined as those rated BB or lower) than for investment-grade bonds, 15.6 basis points versus 12.9 basis points. Trade execution costs dropped by 6.4 basis points post-TRACE for the investment-grade sample. The results also indicate much larger reductions in execution costs (36.0 basis points) after TRACE for non-investment-grade bonds, though the estimate are implausibly large, likely reflecting the small number of low-rated bonds that became TRACE-eligible.

6.3. The effect of TRACE reporting on non-TRACE-eligible bonds

The model presented in Section 3 implies that the public reporting of transactions in a subset of corporate bond issues will result in a liquidity externality of the type described by Amihud et al. (1997), due to improvements in the accuracy of related bond valuation. If so, we anticipate that investors can better evaluate the trade execution costs that they pay in bonds whose transactions are not disseminated through TRACE as well.

Panel A of Table 3 reports the results of estimating expression (18) for the sample of non-TRACE bonds. Not surprisingly, in light of the fact that non-TRACE bonds are of lower average credit quality and trade less frequently, estimated one-way execution costs during the first half of 2002 are considerably greater for non-TRACE bonds (18.3 basis points when estimated by WLS and 20.2 basis points when estimated by GMM) than for the TRACE sample. Most importantly, and consistent with the model presented in Section 3, execution costs for the non-TRACE sample also decreased significantly after the initiation of TRACE reporting. The estimated decline after the initiation of TRACE reporting for the non-TRACE sample is 3.6 basis points (p-value $= 0.015$) or 3.7 basis points (p-value $= 0.042$) based on WLS and GMM estimation, respectively.

Panels B and C of Table 3 provide evidence on cross-sectional variation in the effect of TRACE reporting on non-TRACE bonds. Methods and definitions generally parallel those used for results reported in Table 2 for TRACE-eligible bonds. Several observations can be made. First, in Panel B, Column 1, we find evidence of a decrease in the transaction costs of mid- and high-volume bonds (both are significant at the 10% level), but not for low-volume bonds. Second, in Panel B, Columns 2 and 3, non-TRACE bonds of lower credit ratings also saw greater reductions in trade execution costs, the decrease being 4.9 basis points (p-value $= 0.038$) for non-investment-grade versus an insignificant 2.4 basis point (p-value $= 0.116$) reduction for investment-grade. Third, Panel C Columns 1 and 2, report that trade execution costs for large (over \$500 million original issue value) and small (under \$500 million) bond issues were similar, approximately 20 basis points, pre-TRACE, but only the large issues had significant decrease in transaction costs after TRACE (8.3 basis points, p-value $= 0.001$). The latter result is consistent with the notion that the spillover effect will be stronger for bonds that are more closely related to TRACE-eligible bonds, which are of larger average issue size.

Finally, Table 3, Panel C, Columns 3 and 4 reports results for subsets of the non-TRACE sample, based on the relation between individual bond issues and bonds in the TRACE sample. There is an increase of 4.1 basis points (p-value $= 0.099$) in the cost of trading non-TRACE bonds issued by firms that also have TRACE-eligible bonds. Though the results are only weakly significant, this result is surprising. However, these bonds are smaller (the median bond issue size is \$350 million, compared to \$1.1 billion for TRACE-eligible bonds) and trade less frequently (the median number of trades pre-TRACE is only

Table 3
Spreads on Non-TRACE Bonds
Panel A: Full sample half-spread estimates

In this table, we examine the half-spread on non-TRACE-eligible corporate bonds for 2002 with the basic model described in the Table 2, Panel A heading. In Columns 1 and 3, we examine the half-spread for non-TRACE-eligible bonds in 2002. To assess the impact of TRACE reporting, we interact the ΔQ_{it} variable with a TRACE indicator variable that equals one for trades occurring after June 30, 2002 and zero for earlier trades in Columns 2 and 4. In Columns 1 and 2, we estimate Eq. (2) using weighted least squares (WLS), with the weight being a function of the time between Q and Q_{t-1} and the statistical significance determined from bootstrapped probability estimates. In Columns 3 and 4, we estimate Eqs. (1) and (2) jointly using generalized method of moments (GMM) while controlling for serial correlation (at two lags) and heteroskedasticity.

Estimation technique Column #	WLS (1)	WLS (2)	GMM (3)	GMM (4)
First-stage results				
Intercept	0.0661***	0.0661***	0.1442***	0.1442***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Q_{t-1}	0.1237***	0.1237***	0.2790***	0.2790***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Second-stage results				
Intercept	−0.0657	−0.0654	−0.0285**	−0.0285**
(probability)	(0.151)	(0.150)	(0.013)	(0.013)
Q^*	−0.0275	−0.0276	−0.0347	−0.0352
(probability)	(0.150)	(0.119)	(0.111)	(0.105)
Treasury return	0.4569***	0.4568***	0.4797***	0.4783***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Stock return × investment	0.0557***	0.0558**	0.0527***	0.0529***
(probability)	(0.039)	(0.039)	(0.000)	(0.000)
Stock return × noninvestment	0.0570***	0.0570***	0.0553***	0.0544***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
ΔBaaTrsy × investment	−0.0027**	−0.0027**	−0.0232	0.0001
(probability)	(0.039)	(0.040)	(0.886)	(0.982)
ΔBaaTrsy × noninvestment	−0.0375***	−0.0375***	−0.0367***	−0.0367***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Delta$Q$	0.1825***	0.2023***	0.2024***	0.2226***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Delta$Q$ × TRACE		−0.0357**		−0.0371**
(probability)		(0.015)		(0.042)
Adjusted R^2	9.75%	9.76%	10.79%	10.80%
N	53,237	53,237	53,237	53,237

***, **, *Denote statistical significance at the 99%, 95%, and 90% level, respectively.

Table 3 (*continued*)

Panel B: Half-spreads by firm trade volume and credit rating

In this panel, we examine the impact of TRACE on non-TRACE-eligible bonds' half-spread based on trading volume in Column 1 (bonds are segmented into terciles based on volume), and credit rating in Columns 2 and 3 (if the firm is investment [BBB- and above] or non-investment [below BBB-] grade). The basic model is described in more detail in the Table 2, Panel A heading. We estimate the second stage using weighted least squares (WLS), with the weight being a function of the time between Q and Q_{t-1}. Reported in parentheses are bootstrapped probability values.

	By volume	By credit rating	
Column #	High vol./low vol. (1)	Investment-grade (2)	Non-investment-grade (3)
First-stage results			
Intercept	0.0661***	0.1256***	−0.0520***
(probability)	(0.000)	(0.000)	(0.000)
Q_{t-1}	0.1237***	0.1147***	0.1168***
(probability)	(0.000)	(0.000)	(0.000)
Second-stage results			
Intercept	−0.0649***	−0.0710***	−0.0302
(probability)	(0.000)	(0.000)	(0.167)
Q^*	−0.0276	−0.0506**	0.0132
(probability)	(0.120)	(0.022)	(0.334)
Treasury return	0.4576***	0.6031***	0.1159***
(probability)	(0.000)	(0.000)	(0.000)
Stock return × investment	0.0559***	0.0603***	
(probability)	(0.000)	(0.000)	
Stock return × noninvestment	0.0570***		0.0512***
(probability)	(0.000)		(0.000)
ΔBaaTrsy × investment	−0.0026**	−0.0010***	
(probability)	(0.044)	(0.000)	
ΔBaaTrsy × noninvestment	−0.0376***		−0.0192***
(probability)	(0.000)		(0.000)
Delta$Q$		0.2067***	0.1877***
(probability)		(0.000)	(0.000)
Delta$Q$ × TRACE		−0.0237	−0.0492**
(probability)		(0.116)	(0.038)
Low volume × Delta$Q$	0.2907***		
(probability)	(0.000)		
Low vol. × Delta$Q$ × TRACE	0.0479		
(probability)	(0.261)		
Mid volume × Delta$Q$	0.3218***		
(probability)	(0.000)		
Mid volume × DeltaQ × TRACE	−0.0893*		
(probability)	(0.070)		
High volume × Delta$Q$	0.1739***		
(probability)	(0.000)		
High vol. × Delta$Q$ × TRACE	−0.0226*		
(probability)	(0.068)		
Adjusted R^2	9.82%	13.46%	7.40%
N	53,237	35,728	17,508

***, **, *Denote statistical significance at the 99%, 95%, and 90% level, respectively.

Table 3 (*continued*)

Panel C: Half-spreads by issue size and industry

In this panel, we examine the impact of TRACE on non-TRACE bonds' half-spread based on the size of the bond issue in Columns 1 and 2, if the bond is issued by a firm with TRACE-eligible bonds in Column 3, and if the bond is in the same industry as a TRACE-eligible bond in Column 4. A complete explanation of the basic model is provided in the Table 2, Panel A heading. We estimate the second-stage using weighted least squares (WLS), with the weight being a function of the time between Q and Q_{t-1}. Reported in parentheses are bootstrapped probability values.

By bond characteristics Column #	Bond issues <500 MM (1)	Bond issues $\geqslant 500$ MM (2)	TRACE Firm but non-TRACE bonds (3)	Bonds in the same industry as TRACE bonds (4)
First-stage results				
Intercept	0.0531***	0.0854***	0.0502***	0.0739***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Q_{t-1}	0.1167***	0.1337***	0.1450***	0.0906***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Second-stage results				
Intercept	0.0487***	−0.2089***	−0.2211***	−0.0296**
(probability)	(0.001)	(0.000)	(0.000)	(0.016)
Q^*	−0.0263	−0.0237	0.0853***	−0.0422**
(probability)	(0.215)	(0.221)	(0.005)	(0.048)
Treasury return	0.4282***	0.5176***	0.5314***	0.4510***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Stock return × investment	0.0410**	0.0704**	0.0560***	0.0617***
(probability)	(0.033)	(0.044)	(0.000)	(0.000)
Stock return × noninvestment	0.0494***	0.0884***	0.0767***	0.0581***
(probability)	(0.000)	(0.003)	(0.000)	(0.000)
ΔBaaTrsy × investment	−0.0038	−0.0038	−0.0008	−0.0042**
(probability)	(0.215)	(0.363)	(0.127)	(0.021)
ΔBaaTrsy × noninvestment	−0.0426***	−0.0143***	−0.0184***	−0.0390***
(probability)	(0.000)	(0.000)	(0.008)	(0.000)
Delta$Q$	0.2009***	0.2035***	0.0776***	0.2501***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Delta$Q$ × TRACE	−0.0035	−0.0830***	0.0415*	⁻0.0516***
(probability)	(0.384)	(0.001)	(0.099)	(0.007)
Adjusted R^2	9.70%	10.73%	10.57%	10.48%
N	31,998	21,238	13,220	27,787

***, **, * Denote statistical significance at the 99%, 95%, and 90% level, respectively.

five), and in general we have found that smaller bonds had smaller benefits from the implementation of TRACE. The increase could also reflect the unusual pre-TRACE estimate of trading costs for this subsample, which at 7.8 basis points is the smallest among any subsample reported in Table 2 or 3. Finally, the result could reflect a substitution effect where investors focus their trades on TRACE-eligible bonds issued by the same firm.

Interestingly, we find evidence that other bonds in the same industry (defined by the three-digit SIC code) as an issuer with a TRACE-eligible bond have a significant reduction, 5.2 basis points with an associated p-value of 0.007, in their half-spread after TRACE initiation. For bonds issued by firms outside the industries included in the TRACE sample (not reported) the change in trade execution costs is not statistically significant. This last pair of results is to be expected if observing TRACE trade reports is most useful for improving the pricing and monitoring of trade execution costs for economically similar bonds.

Consistent with the implications of the model developed here, we document a significant liquidity externality, i.e. a decrease in trade execution costs for non-TRACE bonds, after the introduction of transaction reporting. The effect is greatest for larger bonds that have higher volume, and is concentrated in bonds issued by firms in the same industry as TRACE-reported bonds.

6.4. Trade size analysis

The empirical methods used to this point have allowed estimated trade execution costs to vary by subsample, but have not considered possible variation in execution costs as a function of trade size. Trade size can affect percentage execution costs due to economies of scale, differing inventory risk for large trades, or variation in customers' sophistication. Green et al. (2004) and Harris and Piwowar (2005) have each documented substantial variation in trading costs as a function of trade size for municipal bonds, and Edwards et al. (2006) report a similar result for corporate bonds.

We follow Edwards et al. (2006) by allowing the half-spread to be a nonlinear function of trade size (SZ_t):

$$\alpha S_t = b_0 + b_1 * Inverse\,(SZ_t) + b_2^* \operatorname{Log}(SZ_t) + b_3^* SZ_t + b_4^* SZ_t^2 + \ddot{\upsilon}. \tag{20}$$

Expression (20) is substituted into (18), and the resulting expression is estimated by WLS, again using indicator variables to allow coefficient estimates to differ post-TRACE. Table 4 reports fitted values of (20) implied by the resulting coefficient estimates, evaluated at trade sizes of \$0.5 million, \$1 million, \$2 million, \$3 million, \$4 million, \$5 million, and \$10 million. Results for TRACE-eligible bonds are reported in Columns 1–3, and results for non-TRACE bonds are in Columns 4–6. Probability values are based on the block bootstrap method.

For TRACE-eligible bonds, the point estimates indicate that trade execution costs decrease with trade size, which is broadly consistent with the results reported by Green et al. (2004), Harris and Piwowar (2005), and Edwards et al. (2006). The pre-TRACE estimates of execution costs range from 15.1 basis points for trades of \$0.5 million to 9.5 basis points for trades of \$10 million. For non-TRACE bonds during the first half of 2002, the relation between trade size and estimated execution costs is non-monotone. Estimated costs are 22 basis points for trades of \$0.5 million, 23–24 basis points for trade sizes ranging from \$1 to \$5 million, and decrease to 18 basis points for trades of \$10 million.

For both subsamples, the effect of TRACE reporting on transaction costs increases with trade size. We do not detect a significant change in trading costs after TRACE in either sample for transactions of \$1 million or less. For TRACE-eligible bonds, the estimated decrease in trading costs is statistically significant for trades of \$2 million or larger. Point estimates of the decrease in trading costs are 3.4 basis points at trades of \$2 million and 7.5 basis points at trades of \$5 million.[7] For non-TRACE bonds we detect statistically

[7]The point estimates for trades of \$10 million indicate a decrease in costs post-TRACE of 14.7 basis points, to a level of negative 5.2 basis points. This implausibly low estimate of costs likely reflects the paucity of very large trades and the inability of even a relatively complex five-parameter curve such as (20) to perfectly capture the interaction between trading costs and trade size.

Table 4
Trade size analysis
In this table, we report average trading costs for various trade sizes for TRACE-eligible and non-TRACE-eligible bonds, implied by the estimated coefficients of the following two-stage transaction cost model with the first-stage estimated as

$$Q_t = a + bQ_{t-1} + \varepsilon_t. \tag{1}$$

Referring to ε_t from Eq. (1) as Q_t*, the second stage is estimated as

$$\Delta P = a + wX_t + \gamma SQ_t^* + \alpha S\Delta Q + \alpha S\Delta Q^* \text{ TRACE} + \acute{\eta}, \tag{2}$$

where αS is modeled as function of trade size (SZ_t):

$$\alpha S = b_0 + b_1 * \text{ Inverse}\,(SZ_t) + b_2 * \text{Log}\,(SZ_t) + b_3 * SZ_t + b_4 * SZ_t^2 + \ddot{v}. \tag{3}$$

A complete explanation of the basic model is provided in the Table 2, panel A heading. We estimate the second stage using weighted least squares (WLS), with the weight being a function of the time between Q and Q_{t-1}. The estimated costs for a trade of size SZ are those implied by the coefficients of the two-stage transaction cost model. In Columns 1, 2, and 3, we report estimated trade costs for various trade sizes for TRACE-eligible bonds. In Columns 4, 5, and 6 we report estimated trade costs for various trade sizes for non-TRACE-eligible bonds. Reported in parentheses are bootstrapped probability values.

	TRACE bonds			Non-TRACE bonds		
Trade Size ($1,000)	Before TRACE (1)	After TRACE (2)	Impact of TRACE (3)	Before TRACE (4)	After TRACE (5)	Impact of TRACE (6)
500	0.1513***	0.1330***	−0.0183	0.2186***	0.2232***	0.0046
(probability)	(0.000)	(0.000)	(0.195)	(0.000)	(0.000)	(0.394)
1,000	0.1358***	0.1137***	−0.0221	0.2331***	0.2372***	0.0040
(probability)	(0.000)	(0.000)	(0.146)	(0.000)	(0.000)	(0.407)
2,000	0.1210***	0.0873*	−0.0337***	0.2417***	0.2368***	−0.0048
(probability)	(0.000)	(0.067)	(0.000)	(0.000)	(0.000)	(0.408)
3,000	0.1130***	0.0661***	−0.0469**	0.2411***	0.2250***	−0.0161
(probability)	(0.000)	(0.000)	(0.026)	(0.000)	(0.000)	(0.186)
4,000	0.1079***	0.0471***	−0.0608***	0.2367***	0.2088***	−0.0279*
(probability)	(0.000)	(0.007)	(0.007)	(0.000)	(0.007)	(0.061)
5,000	0.1042***	0.0293*	−0.0750***	0.2301***	0.1904***	−0.0398**
(probability)	(0.000)	(0.069)	(0.000)	(0.000)	(0.000)	(0.017)
10,000	0.0953***	−0.0517**	−0.1469***	0.1820***	0.0867***	−0.0953***
(probability)	(0.000)	(0.019)	(0.000)	(0.000)	(0.000)	(0.001)

***, **, * Denote statistical significance at the 99%, 95%, and 90% level, respectively.

significant decreases in trading costs only for trades of $4 million and larger, with point estimates of the decrease being 2.8 basis points at trades of $4 million, 4.0 basis points at trades of $5 million, and 9.5 basis points at trades of $10 million.

To summarize, the estimates obtained thus far indicate that trade execution costs were reduced significantly (by about 50% for TRACE-eligible and by about 20% for non-TRACE eligible bonds) after market transparency was improved due to TRACE during 2002. Finding large reductions in trade execution costs in the present sample of institutional trades, and particularly finding greater reductions for larger trades, is not expected if one conjectures that a lack of transparency is mainly a problem for unsophisticated small traders. Rather, these results indicate that even the market for large institutional corporate bond trades has been substantially affected by transaction reporting.

6.5. Robustness tests

The empirical results reported in the preceding sections indicate that estimated trade execution costs for insurance company trades in corporate bonds decreased in the second half of 2002 as compared to the first half, both for bonds whose trades were disseminated through TRACE and for bonds whose trades were not disseminated through TRACE. We next report the results of a series of robustness tests.

6.5.1. Alternative cost measures

Some studies, e.g., Chakravarty and Sarkar (2003), Hong and Warga (2000), and Goldstein et al. (2005), have estimated "traded spreads" by comparing the average dealer selling price to the average dealer purchase price when both purchases and sales are observed for the same bond within a reasonably short time interval (e.g., the same day). Methods that rely on matched dealer purchases and sales provide estimates of trading costs when trades can be paired, but require that many transactions be ignored when dealer purchases cannot be well matched with comparable dealer sales. This shortcoming will be particularly pertinent for bonds that are not frequently traded. However, we report estimates of transactions cost obtained using this methodology, for robustness.

We construct a subsample of trades consisting of buys and sells of the same bond on the same day, which reduces the TRACE (non-TRACE) sample from 39,040 (53,282) trades to 9,567 (9,543) trades. In cases with more than one buy (sell) trade during a day, we take the simple average of prices. This leaves 2,296 (2,572) matched pairs of buy and sell prices for the same bond on the same day.

For TRACE-eligible bonds the estimated percentage traded half-spread is 14.7 basis points pre-TRACE and 10.0 basis points post-TRACE. For the non-TRACE sample the estimated percentage half-spread is 14.1 basis points before TRACE and 9.2 basis points after TRACE implementation. Thus, this method also supports the conclusion that execution costs decreased post-TRACE. However, the estimated decline in traded spreads obtained by this method is not statistically significant, either for the TRACE (p-value $=$ 0.230) or non-TRACE sample (p-value $= 0.110$). The lack of statistical significance mainly reflects the decreased sample size.

As an additional robustness test, we estimate the basic model (18) using a six-month window centered on the introduction of TRACE. Based on this shorter horizon, the estimated reductions in trading costs are 8.0 basis points (p-value < 0.000) for TRACE-eligible securities and 2.3 basis points (p-value $= 0.167$) for non-TRACE securities. The decreased statistical power inherent in the use of this short window precludes the ability to identify significant subsample results.

6.5.2. Controlling for variation in the economic environment

We choose to attribute the large decreases in estimated trading costs (50% for TRACE-eligible bonds and 20% for non-TRACE-eligible bonds) after TRACE introduction to the improved ability to monitor and control trade execution costs in the more transparent environment. An alternative view, however, is that some or all of the reduction in trade execution costs is attributable to changes in the economic environment other than the introduction of TRACE reporting.

To assess this possibility we expand expression (18) to include variables that could potentially affect bid-ask spreads in corporate bond markets. Suppose that the spread for trade t depends on variable Z according to

$$\alpha S_t = b_0 + b_1\ \text{TRACE} + b_2 Z'_t, \tag{21}$$

where the $'$ denotes that variable Z is expressed as deviations from its own time-series mean. Substituting (21) into (18) gives an expanded indicator regression model:

$$\Delta P_t = a + wX + \gamma Q_t^* + b_0 \Delta Q + b_1\ \text{TRACE}^* \Delta Q + b_2 Z'_t \Delta Q + \dot{\eta}. \tag{22}$$

In expression (22), the coefficient b_0 estimates trading costs pre-TRACE, the coefficient b_1 estimates the effect of TRACE reporting on trading costs conditional on a specific (i.e., the time-series mean) outcome on the explanatory variable Z, while the coefficient b_2 estimates the effect of variable Z on the half-spread. Candidates for inclusion in Z should be variables that plausibly affect the costs of corporate bond market making. Numerous authors, beginning with Demsetz (1968), have argued that increased trading volume should reduce bid-ask spreads. We accordingly include a measure of trading activity. Given that many individual bonds trade infrequently and that returns across various bonds are likely to be highly correlated, we employ a simple market-wide measure of trading activity, the total number of trades contained in the sample over the prior five trading days. Other authors, at least since Ho and Stoll (1980), have emphasized that dealers will widen spreads if inventory risk increases. To allow for possible effects of interest rate risk, we obtain the one-day-ahead conditional variance estimate from a GARCH(1,1) model applied to the time series of returns to the 10-year on-the-run Treasury note.

Results of estimating expression (22) are reported in Table 5, Columns 1 and 2, for TRACE-eligible and non-TRACE bonds, respectively. The results do not support the reasoning that interest rate volatility is a determinant of spreads for corporate bonds in the present sample, as the coefficient estimate on the product of ΔQ and the interest rate volatility measure is not significant (p-values are 0.362 and 0.351, respectively) in either TRACE-eligible or non-TRACE-eligible bonds. The results do support the reasoning that bid-ask spreads for corporate bonds decrease with trading activity, as point estimates on the product of ΔQ and recent trading activity are negative and significant, particularly for TRACE bonds.

Most importantly, the results indicate that allowing for the possible effects of changes in trading volume and interest rate volatility on spreads does not alter the key conclusion that spreads decreased significantly after the introduction of TRACE. The estimated effect of TRACE reporting (b_1) for TRACE-eligible bonds is 10.8 basis points (p-value $= 0.001$), which is larger than the 7.9 basis points in Table 2. The corresponding estimate for non-TRACE bonds is 3.9 basis points (compared to 3.6 basis points in Table 3) and is still significant (p-value $= 0.043$). We conclude that changes in interest rate volatility and trading activity do not explain the reduction in corporate bond bid-ask spreads post-TRACE.

Table 5
Robustness

Panel A. Additional economic controls and dealer analysis

In this table, we include additional control variables and analyze the impact of TRACE by dealer size. In Columns 1 and 2, we include measures that could be influencing the spread on corporate bonds, namely the volatility of the bond market and prior trading volume. Volatility is estimated as the conditional heteroskedasticity predicted by estimating a $GARCH_{(1,1)}$ model of the return on the ten year on-the-run Treasury note. Trading volume is measured as the summation of the prior five-day trading volume (number of trades). These variables are then interacted with the spread measure (Delta$Q$). In Columns 3 and 4, we examine if the impact of TRACE differed based on the size of the dealer completing the transaction. To do this, we first segment trades into those facilitated by large dealers (12 large dealers are identified) or small dealers and create an indicator variable. We then interact this variable with Delta$Q$ and Delta$Q \times$ TRACE to provide estimates for the differential of trading costs between large and small dealers pre- and post-TRACE. A complete explanation of the basic model is provided in the heading of Table 2, Panel A. We estimate the second stage using weighted least squares (WLS), with the weight being a function of the time between Q and Q_{t-1}. Reported in parentheses are bootstrapped probability values.

	Economic controls		Dealer analysis	
Column #	TRACE (1)	Non-TRACE (2)	TRACE (3)	Non-TRACE (4)
First-stage results				
Intercept	0.0884***	0.0661***	0.0884***	0.0661***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Q_{t-1}	0.0785***	0.1237***	0.0785***	0.1237***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Second-stage results				
Intercept	0.0075**	−0.0668	−0.0091***	−0.0681
(probability)	(0.023)	(0.155)	(0.001)	(0.317)
Q^*	−0.0165	0.0273**	−0.0127	−0.0269
(probability)	(0.321)	(0.021)	(0.395)	(0.130)
Treasury return	0.2232***	0.4566***	0.2207***	0.4564***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Stock return × investment	0.0480***	0.0560***	0.0500***	0.0560***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Stock return × noninvestment	0.0801***	0.0583***	0.0797***	0.0570***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
ΔBaaTrsy × investment	−0.0014	−0.0025**	−0.0004	−0.0027**
(probability)	(0.119)	(0.048)	(0.222)	(0.040)
ΔBaaTrsy × noninvestment	−0.0446***	−0.0375***	−0.0449***	−0.0374***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Delta$Q$	0.1926***	0.2020***	0.1795***	0.2258***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Delta$Q$ × large dealer			−0.0746***	−0.0425**
(probability)			(0.005)	(0.011)
Delta$Q$ × TRACE	−0.1077***	−0.0386**	−0.1049***	0.0041
(probability)	(0.001)	(0.043)	(0.001)	(0.414)
Delta$Q$ × TRACE × large dealer			0.0403	−0.0758**
(probability)			(0.122)	(0.000)
Delta$Q$ × $TRSYGARCH_{(1,1)}$	0.1115	0.0645		
(probability)	(0.362)	(0.351)		
Delta$Q$ × Trading Volume	−0.0035**	−0.0023***		
(probability)	(0.017)	(0.000)		
Adjusted R^2	4.89%	9.91%	3.25%	9.82%
N	38,970	53,237	37,645	50,854

***, **, * Denote statistical significance at the 99%, 95%, and 90% level, respectively.

Table 5 (*continued*)

Panel B: Trend analysis & expansion of TRACE-eligible bonds on March 3, 2003

To control for a time trend, we examine bonds that trade during both 2001 and post-TRACE in 2002. We separate these bonds into groups that did or did not become TRACE-eligible on July 1, 2001 and then reexamine the basic specification while including a daily time trend variable in Columns 1 and 2. On March 3, 2003 TRACE was expanded to cover all bonds rated A or above with an issue size of at least $100 million. This group was expanded on April 14, 2003 to include 120 representative BBB bonds. We create a TRACE 2003 sample, that includes all bonds that became eligible on either March 3 or April 14. Our non-TRACE 2003 sample includes all bonds that are not TRACE-eligible in 2003. Similar to our prior analyses, we use a six month pre and post period. However, we exclude all trades between March 1 and April 14 for both samples. In the third Column, we examine the half-spread for bonds that became TRACE-eligible in 2003. In the fourth Column, we examine the half-spread for bonds that were not TRACE eligible in 2003. We estimate the second Eq. using weighted least squares (WLS), with the weight being a function of the time between Q and Q_{t-1}. Reported in parentheses are bootstrapped probability values.

	1/01/01 Thru 12/31/02		Second round TRACE 2003	
	TRACE 2002	Non-TRACE	TRACE 2003	Non-TRACE
	(1)	(2)	(3)	(4)
First-stage results				
Intercept	0.1044***	0.0790***	0.2188***	0.0526***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Q_{t-1}	0.0792***	0.1183***	0.0764***	0.0788***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Second-stage results				
Intercept	−0.0187	−0.0099	0.0700	0.2719
(probability)	(0.443)	(0.483)	(0.151)	(0.150)
Q^*	−0.0330**	−0.0538***	−0.0363**	−0.0257***
(probability)	(0.048)	(0.001)	(0.050)	(0.119)
Treasury return	0.2014***	0.3089***	0.6923***	0.3391***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Stock return × investment	0.0501***	0.0381***	0.0190***	0.0571**
(probability)	(0.000)	(0.000)	(0.000)	(0.039)
Stock return × noninvestment	0.0470***	0.0528***	0.0317***	0.0212***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
ΔBaaTrsy × investment	−0.0005	0.0003	0.0115**	0.0081**
(probability)	(0.220)	(0.153)	(0.039)	(0.040)
ΔBaaTrsy × noninvestment	−0.0368***	−0.0318***	0.0125***	0.0070***
(probability)	(0.000)	(0.000)	(0.010)	(0.000)
Delta$Q$	0.1322***	0.2046***	0.1151***	0.1354***
(probability)	(0.000)	(0.000)	(0.000)	(0.000)
Delta$Q$ × TRACE	−0.0672*	−0.0438**	−0.0323***	−0.0465**
(probability)	(0.057)	(0.039)	(0.003)	(0.015)
Delta$Q$ × TREND	0.0000	0.0001*		
(probability)	(0.478)	(0.080)		
Adjusted R^2	3.25%	7.50%	24.18%	8.82%
N	49,499	68,353	21,563	27,631

***, **, * Denote statistical significance at the 99%, 95%, and 90% level respectively.

6.6. Dealer market shares

We next examine whether the execution costs of corporate bond trades differ across large and small dealers, and whether the introduction of transaction reporting has altered the competitive structure of the bond market. The NAIC database contains the name of the specific dealer involved in a trade. Based on those trades completed before TRACE implementation, we classify the 12 largest dealers in the database as large dealers, and classify the remaining dealers as small dealers. However, dealer information is missing for some trades. Omitting trades that lack dealer information reduces the sample to 37,645 trades for TRACE-eligible bonds and 50,854 trades for non-TRACE bonds.

We then calculate a 12-dealer concentration ratio before and after TRACE introduction, and following Schultz (2001), define an indicator variable (*LargeDealer)* that equals one for trades facilitated by one of the 12 large dealers and zero otherwise. For each bond issue, dealer market share is computed as total dollar trading volume by dealer divided by the cumulative volume for the bond issue. The concentration ratio (CR) is the sum of the market share for the 12 largest dealers in the bond, where market share is calculated separately before and after TRACE introduction for each bond. For TRACE-eligible bonds, we estimate an average CR of 55.6% pre-TRACE and 45.4% post-TRACE (p-value of difference $= 0.000$), indicating a reduction in the market share of large dealers. For non-TRACE bonds, the CR is 51.7% pre-TRACE and declines to 45.1% post-TRACE (p-value of difference $= 0.001$). The decline in concentration of trading is consistent with the reasoning that the market has become more competitive after TRACE introduction.

Columns 3 and 4 of Table 5 report results of estimating a version of (18) that allows for differential trading costs when transacting with large dealers. We include in the regression the product of the indicator variable *LargeDealer* and ΔQ, with and without the post-TRACE indicator. For TRACE-eligible bonds, the coefficient on the interaction term pre-TRACE is -0.0746 (p-value $= 0.005$). This implies lower trading costs when transacting with large dealers, which is consistent with Schultz's (2001) finding. More importantly, this effect is diminished in the post-TRACE period, as the coefficient on the interaction term is -0.033 ($-0.075+0.043$), which is statistically insignificant (p-value $= 0.242$). This result is consistent with the reasoning that the corporate bond market has become more competitive after TRACE introduction. Large dealers were able to charge lower trade execution costs before TRACE introduction, perhaps because of informational advantages inherent in observing a larger proportion of order flow. The advantage of observing order flow was reduced after TRACE introduction, as even dealers that do not participate in trades can observe prices.

For the non-TRACE bonds, the coefficient of the interaction term pre-TRACE is -0.043 (p-value $= 0.011$), and is -0.119 (-0.043–0.076) post-TRACE (p-value $= 0.000$). These results indicate that for non-TRACE bonds, large dealers continue to charge less than small dealers even after TRACE initiation, suggesting that trading costs could decline further when transaction reporting commences for these ineligible bonds.

6.7. Time trend

If trading costs were trending downward through time, the decrease in costs during the second half of 2002 could simply reflect the continuation of this trend rather than the introduction of TRACE reporting. We consider this result unlikely, both because of the magnitude of the decrease in estimated costs during the second half of the year, and also because our trading cost estimates from the first half of 2002 are generally similar to the estimates provided by Schultz (2001) for 1995–1997.

To further evaluate the possibility of a time trend in costs, we construct a sample of bonds that traded at least once during each half-year from January 1, 2001 to December 31, 2002. We then estimate the basic model (18) while including a trend variable: the number of elapsed trading days since January 1, 2002. The coefficient on this variable will capture any linear time trend in trading costs.

The results of this analysis are reported in Columns 1 and 2 of Panel B in Table 5. We find the time-trend variable to be insignificant in the TRACE sample (point estimate 0.0000, p-value $= 0.478$), and positive and weakly significant (point estimate 0.0001, p-value $= 0.080$) in the non-TRACE sample, indicating a very slight upward, rather than downward, trend in trading costs.

Estimated trading costs for the January 2001–June 2002 period are 13.2 basis points for TRACE bonds (Column 1) and 20.5 basis points for non-TRACE bonds, which are very similar to the corresponding estimates of 13.5 basis points (Column 2 of Panel A in Table 2) and 20.5 basis points (Column 2 of Panel A in Table 3) reported previously for the first half of 2002. Most importantly, inclusion of the linear time trend variable does not alter inference with respect to TRACE introduction. For the TRACE sample, the estimated impact of TRACE is -6.7 basis points, compared to -7.9 basis points in Table 2. For the non-TRACE sample, the estimated impact of TRACE is -4.4 basis points, versus the estimated -3.6 basis points in Table 3. While not reported, we also estimate the model with a nonlinear time-trend specification and find similar results. Thus, we find no evidence that the key results reported here are driven by a secular time trend in trading costs.

6.8. Expansion of TRACE-eligible bonds in 2003

There were two additional significant expansions during 2003 in the set of bonds eligible for TRACE transaction reporting.[8] On March 3, 2003 TRACE was expanded to cover all bonds rated A and higher with an issue size of at least \$100 million. On April 14, 2003 an additional 120 representative BBB bonds became eligible for TRACE transaction reporting. We create a TRACE 2003 sample, which includes all bonds that became eligible on either March 3 or April 14. Our non-TRACE 2003 sample includes all bonds that remained non-TRACE-eligible throughout 2003. To correspond to our prior analysis, we estimate trade execution costs for the six-month period before and after the initiation of TRACE reporting. However, since the two events are closely clustered in time, we define the pre-TRACE period as the six months before March 3, 2003 and the post period as the six months after April 14, 2003, omitting trades between the two dates.

The results of estimating (18) for the newly TRACE-eligible bonds are reported in Column 3, Panel B, of Table 5, while results for bonds that remained non-TRACE-eligible through 2003 are found in Column 4. For the newly TRACE-eligible bonds, the pre-TRACE estimate of trading costs is 11.5 basis points, and the decline in trading costs after TRACE introduction is 3.2 basis points (p-value $= 0.003$).

For the bonds that remained non-TRACE-eligible, the estimated half-spread is 13.5 basis points prior to the expansion in the set of bonds eligible for TRACE transaction reporting, and declines by 4.7 basis points after April 14, 2003. Thus, these results support

[8] There were also two more minor changes in 2003 as some bonds entered and an equivalent number were deleted from the set of 50 non-investment-grade bonds subject to TRACE transaction reporting.

similar conclusions as our earlier analysis: estimated trade execution costs decreased after the commencement of trade reporting, both for those bonds whose trades are disseminated through TRACE and for bonds whose trades are not yet disseminated through TRACE. Our results regarding the 2003 TRACE expansions are also broadly consistent with those reported by Goldstein et al. (2005) and Edwards et al. (2006), who examine the 2003 corporate bond market in more detail.

7. Conclusion

We develop a theoretical model and test its implications by estimating trade execution costs for a sample of institutional (insurance company) trades in corporate bonds before and after the initiation of public transaction reporting through the TRACE system for a subset of bonds in July 2002. The results indicate a remarkable (approximately 50%) reduction in trade execution costs for bonds eligible for TRACE transaction reporting. In addition, trade execution costs for bonds not eligible for TRACE reporting decreased by about 20%, which likely reflects a liquidity externality by which better pricing information regarding a subset of bonds improves valuation and execution cost monitoring for related bonds. The cumulative dollar impact of these trading cost reductions are large – extrapolating beyond the present sample, a rough "back of the envelope" calculation suggests annual trading cost reductions of about $1 billion for the full corporate bond market.[9] This estimate is no doubt imprecise, as it relies on simplifying assumptions, including that the experience of the insurance companies comprising the present sample is representative of the broader market. However, the magnitude of the estimate emphasizes the potential economic importance of designing market mechanisms optimally.

The results reported here indicate decreased trade execution costs for relatively large institutional traders after the implementation of TRACE. Madhavan (1995) develops a model implying that trade disclosure will reduce transaction costs for small noise traders, but increase costs for large traders. His model assumes fully competitive markets with zero average dealer profits in each setting, and no change in the costs of market making. Our finding that trading costs are reduced for large institutional traders in the more transparent market is therefore consistent with the reasoning that market makers earned economic rents in the opaque market, or that the costs of market making are lower in the more transparent environment. Further research to distinguish between these possibilities would be useful. It would also be productive to assess how TRACE reporting and the associated reduction in trading costs has altered the behavior of investors, market makers, and issuing firms. For example, a reduction in the benefits to superior information could adversely affect incentives for traders to incur costs in order to become informed, which could in turn affect the informational efficiency of the bond markets. Further, improved liquidity in

[9]The post-TRACE sample contains $64.5 billion in trading in TRACE bonds and $83.9 billion in trading in non-TRACE bonds. We estimate trading cost reductions ranging from 4.9 to 7.9 basis points for the TRACE sample (Table 2, Panel A) and 3.6 to 3.7 basis points (Table 3, Panel A) for the non-TRACE sample. Using the lower estimates for conservatism, this equates to $31.6 million and $30.2 million in trading cost reductions for the TRACE and non-TRACE samples, respectively, or a total of $61.8 million for our sample of insurance company trades during the second half of 2002. As noted in Section 4.2. insurance companies participated in 12.5% of the dollar volume for TRACE-eligible bonds during the second half of 2002. If this proportion is representative, the annualized equivalent estimate for the full market is $(61.8/.125) \times 2 =$ $989 million.

corporate bond markets could alter firms' optimal capital structures, as implied by the model of Chang and Yu (2004).

The results reported here are important because they verify that market design, and in particular decisions as to whether to make the market transparent to the public, have first-order effects on the costs that customers pay to complete trades. Further, since the sample employed here consists of institutional trades, these results indicate that public trade reporting is important not only to relatively unsophisticated small traders, but also to professional investors who make multi-million dollar transactions.

We document large reductions in trade execution costs after transaction reporting in corporate bonds was initiated. However, we caution that this analysis does not yet identify the optimal degree of transparency. Increased transparency could affect incentives to expend resources to become informed about non-public information that affects values, which could alter the informational efficiency of the market. Further, relations between market liquidity and transparency might not be linear, implying that further increases in transparency (e.g., the dissemination of quotes and/or limit order information) might or might not be accompanied by further improvements in liquidity. Additional investigation of relations between market transparency, liquidity, and informational efficiency is warranted.

Appendix A:. Simulation evidence on potential bias in coefficients estimated without time stamps

To assess whether unbiased coefficient estimates can be obtained when implementing indicator variable regressions in datasets that do not contain reliable time stamps, we create simulated data where the underlying parameters are known, and then examine the coefficient estimates obtained from applying Eq. (18) while using various techniques with regard to time ordering of the data. The simulated data are created as follows. The trade indicator data series Q is created as a binomial random variable that takes the values 1 and -1 with equal probability, implying that the serial correlation in order flow, ρ, equals zero. The observable and unobservable public information variables X and U are created as normal random variables with mean zero and standard deviation 5.0. Trade prices are specified as in (1a). Bond value is set to an initial value of $V_0 = \$1{,}000$, and then evolves according to (14). Each random variable is independent of the others. The parameters of the model are set to $w = 0.5$ (the weighting on observable public information), $\gamma S = \$1.00$ (the price impact of trades), and $(1—\gamma)S = \$0.50$ (the non-informational component of the half-spread). After a simulated data series is created, the parameters of regression specification (18) are estimated and saved. The entire simulation is repeated one thousand times, creating a distribution of parameter estimates.

Table A.1 reports the mean and standard deviation of the parameter estimates obtained from estimating versions of (18) in the simulated data, when the number of trades in any given simulation is varied from 250 to 100,000. For results reported in Panel A the true time ordering of the data is preserved, so that each change is calculated from the observation that actually preceded it. Expression (18) is the correct specification for the time-ordered data, so we focus primarily on the effect of omitting variables from the regression specification. For the results reported in Panels B and C we divide the simulated data series into a large number of "days," and randomly reorder the observations within each day, while maintaining proper time ordering across days. A trading day is ended when

Table A.1
Regression estimates obtained in simulated data

The simulated data are created by initializing $V_0 = \$1{,}000$, and allowing expectations of value to evolve according to text expression (14). Trade prices are as specified by text expression (1a), with the trade indicator variable Q_t taking the values 1 and -1 with equal probability. Observable and non-observable public information variables X_t and μ_t are simulated as independent normal variables with mean 0 and standard deviation 5.0. The key parameters are $S = \$1.50$, $\gamma S = 1.00$, and $w = .50$. The regression specification is

$$\Delta P = a + wX_t + \gamma SQ_t + (1-\gamma)S\Delta Q + \omega_t$$

which is text Eq. (18) in the case of no autocorrelation in the trade indicator variable. The simulation is repeated 1,000 times. Coefficient and SD denote the mean and standard deviation of the 1,000 coefficient estimates. Estimated intercepts are not reported. For results reported in Panel A, the data are correctly ordered, but variables are omitted from the second two sets of results. For results reported in Panel B, the data are randomly ordered within a "day" and changes are computed with respect to the prior observation. For results reported in Panel C, data are randomly ordered within a "day," but changes are computed with respect to the last trade on the prior "day."

Number trades each simulation	$N = 250$		$N = 1000$		$N = 10{,}000$		$N = 100{,}000$	
	Coefficient	SD	Coefficient	SD	Coefficient	SD	Coefficient	SD
Panel A: Estimation with correctly ordered data								
W (true = 0.5)	0.500	0.031	0.500	0.016	0.500	0.005	0.500	0.002
γS (true = 1.00)	1.004	0.228	1.004	0.110	1.000	0.035	1.000	0.011
(1-γ)S (true = 0.50)	0.497	0.158	0.498	0.079	0.500	0.024	0.500	0.008
γS (true = 1.00)	0.999	0.324	1.001	0.160	1.000	0.049	1.000	0.016
(1-γ)S (true = 0.50)	0.499	0.229	0.498	0.111	0.500	0.034	0.500	0.011
(1-γ)S (true = 0.50)	0.999	0.159	0.999	0.078	0.999	0.025	1.000	0.008
Panel B: Estimation with data randomly ordered within each "day"								
W (true = 0.5)	0.499	0.074	0.500	0.037	0.500	0.011	0.500	0.004
γS (true = 1.00)	0.374	0.302	0.325	0.153	0.320	0.049	0.329	0.014
(1-γ)S (true = 0.50)	0.811	0.269	0.838	0.131	0.839	0.044	0.835	0.013
Panel C: Estimation with data where changes are computed based on prior-day reference trade								
w (true = 0.5)	0.498	0.085	0.499	0.045	0.500	0.015	0.500	0.005
γS (true = 1.00)	0.982	0.831	0.999	0.415	0.995	0.134	1.000	0.040
(1-γ)S (true = 0.50)	0.482	0.769	0.493	0.379	0.500	0.119	0.500	0.036

a "count" variable that is set to one at the beginning of the day, and that either increments by one after each trade or remains unchanged with equal probability, reaches a total of three. The number of trades per day averages three, but can be as low as one or, on rare occasions, over 20. For results reported in Panel B we ignore the fact that trades are not properly time ordered, and simply compute each change as the difference between the observation and the one immediately preceding it in the dataset.

Several results reported in Panel A of Table A.1 are worth noting. First, as expected, the estimation of specification (18) in correctly ordered data provides coefficient estimates that are unbiased and whose standard deviations decline rapidly as the simulated sample size increases. Second, omitting the public information variable X from the regression does not bias coefficient estimates, but does increase the standard deviation of the estimates. For example, with $N = 1{,}000$ simulated trades, omitting the public information variable X increases the standard deviation of the coefficient estimate on ΔQ by 44%. As might be expected, this effect is greater if the simulation is repeated with a larger standard deviation of the simulated public information variable. Third, omitting the trade indicator variable from the regression leads to bias in the estimate of the half-spread.[10] This result can be interpreted simply as the effect of omitting a correlated variable. Since Q can take only two values, its changes are highly correlated with its level. Omitting Q biases the coefficient estimate on the change in Q, unless the true parameter on lagged Q is zero. Even with 100,000 simulated observations, the average estimate of ΔQ obtained from the improper specification that omits Q_t is 1.000, which differs from the true half-spread of 0.500 by more than 60 standard deviations.

Two methodological conclusions are supported by the simulation results reported in Panel A. First, it is desirable to include measures of changes in public information when using an indicator variable regression to estimates spreads, in any situation where changes in the public information set between trades are large relative to spreads. This is likely to be particularly true for bonds and other assets that are traded infrequently. Second, it is important to include the indicator variable Q_t in the specification, unless there is strong reason to believe that the associated parameter, γ, is zero.

Results reported in Panel B of Table A.1 indicate that simply computing changes from the prior observation in data that are not properly time ordered will lead to biased and inconsistent coefficient estimates. Even with $N = 100{,}000$ simulated trades for each round of the simulation, the mean estimates of the non-informational and informational components of the half-spread are 0.835 and 0.329, respectively, which differ from the true parameter values of 0.50 and 1.00 by 26 and 48 standard deviations, respectively.

Panel C of Table A.1 investigates the properties of an alternate estimation strategy that allows for improper time ordering within a day, while still taking advantage of known transaction dates. For each observation, changes are computed as the current observation minus the last observation contained in the dataset (which is not necessarily chronologically last) on the prior trading day. Also for Panel C we redefine the public information variable X as the change in the accumulated change since the last observation on a prior trading day.

The most important result to emerge from this simulation exercise is that the estimation procedure used for the results reported in Panel C of Table A.1 leads to unbiased coefficient estimates. The average parameter estimates are always close to the true parameter values, and the standard deviation of the estimates decreases rapidly as the number of simulated trades increases. This estimation technique could also prove useful in other cases where datasets do not contain time stamps, or where time stamps are not fully reliable.

However, as would be expected given that some information is lost due to the lack of proper time ordering, the standard deviations of the estimates are generally three to five times larger than the standard deviations of the estimates obtained in correctly ordered data as reported in Panel A. The lack of time ordering in the data can therefore be expected to reduce statistical power.

[10] This result can be interpreted simply as the effect of omitting a correlated variable. Since Q can take only two values, its changes are highly correlated with its level. Omitting Q biases the coefficient estimate on the change in Q, unless the true parameter on lagged Q is zero.

References

Amihud, Y., Mendelson, H., Lauterbach, B., 1997. Market microstructure and securities values: evidence from the Tel Aviv Stock Exchange. Journal of Financial Economics 45, 365–390.

Bessembinder, H., Venkataraman, K., 2004. Does an electronic stock exchange need an upstairs market. Journal of Financial Economics 73, 3–36.

Bloomfield, R., O'Hara, M., 1999. Market transparency: who wins and who loses. Review of Financial Studies 12, 5–35.

Bloomfield, R., O'Hara, M., 2000. Can transparent markets survive. Journal of Financial Economics 55, 425–459.

Boehmer, E., Saar, G., Yu, L., 2005. Lifting the veil: an analysis of pre-trade transparency at the NYSE. Journal of Finance 60, 783–815.

Campbell, J., Taksler, G., 2003. Equity volatility and corporate bond yields. Journal of Finance 58, 2321–2349.

Carlstein, E., 1986. The use of subseries methods for estimating the variance of a general statistic from a stationary time series. Annals of Statistics 14, 1171–1179.

Chakravarty, S., Sarkar, A., 2003. Trading costs in three U.S. bond markets. Journal of Fixed Income 13, 39–48.

Chang, C., Yu, X., 2004. Informational efficiency and liquidity premium at the determinants of capital structure. Working paper, Indiana University.

Chen, L., Lesmond, D., Wei, J., 2003. Bond liquidity estimation and the liquidity effect in yield spreads. Working paper, University of Toronto.

Demsetz, H., 1968. The cost of transacting. The Quarterly Journal of Economics 82, 33–53.

Edwards, A., Harris, L., Piwowar, M., 2006. Corporate bond market transparency and transactions costs. Journal of Finance, forthcoming.

Efron, B., Tibshirani, R., 1993. An Introduction to the Bootstrap. Chapman & Hall, London.

Flood, M.R., Huisman, Koedijk, K., Mahieu, R., 1999. Quote disclosure and price discovery in multiple-dealer financial markets. Review of Financial Studies 12, 37–59.

Gemmill, G., 1996. Transparency and liquidity: a study of block trades on the London Stock Exchange under different publication rules. Journal of Finance 51, 1765–1790.

Goldstein, M., Hotchkiss, E., Sirri, E., 2005. Transparency and liquidity: a controlled experiment on corporate bonds. Working paper, Boston College.

Green, R., Hollifield, B., Schurhoff, N., 2004. Financial intermediation and the costs of trading in an opaque market. Working paper, Carnegie Mellon University.

Hall, P., Jing, B., 1996. On sample reuse methods for dependent data. Journal of the Royal Statistical Society, Series B 58, 727–737.

Harris, L., Piwowar, M., 2005, Secondary trading costs in the municipal bond market. Journal of Finance. forthcoming.

Hasbrouck, J., 1993. Assessing the quality of a security market: a new approach to transaction cost measurement. Review of Financial Studies 6, 191–212.

Ho, T., Stoll, H., 1980. On dealer markets under competition. Journal of Finance 35, 259–267.

Hotchkiss, E., Ronen, T., 2002. The informational efficiency of the corporate bond market: an intraday analysis. Review of Financial Studies 15, 1325–1354.

Hong, G., Warga, A., 2000. An empirical study of corporate bond market transactions. Financial Analysts Journal 56, 32–46.

Huang, R., Stoll, H., 1997. The components of the bid-ask spread: a general approach. Review of Financial Studies 10, 995–1034.

Krishnan, C., Ritchken, P., Thomson, J., 2004. Monitoring and controlling bank risk: does risky debt help. Journal of Finance 59, 343–378.

Lesmond, D., Ogden, J., Trzcinka, C., 1999. A new estimate of transactions costs. Review of Financial Studies 12, 1113–1141.

Madhavan, A., 1995. Consolidation, fragmentation, and the disclosure of trading information. Review of Financial Studies 8, 579–603.

Madhavan, A., Cheng, M., 1997. In search of liquidity: block trades in the upstairs and downstairs markets. Review of Financial Studies 10, 175–203.

Madhavan, A., Porter, D., Weaver, D., 2005. Should securities markets be transparent. Journal of Financial Markets 8, 265–287.

Madhavan, A., Richardson, M., Roomans, M., 1997. Why do securities prices change: a transactions level analysis of NYSE-listed stocks. Review of Financial Studies 10, 1035–1064.

Naik, N., Nueberger, A., Viswanathan, S., 1999. Disclosure regulation in markets with negotiated trades. Review of Financial Studies 12, 873–900.

Pagano, M., Roell, A., 1996. Transparency and liquidity: a comparison of auction and dealer markets with informed trading''. Journal of Finance 51, 579–611.

Schultz, P., 2001. Corporate bond trading costs: a peek behind the curtain. Journal of Finance 56, 677–698.

Warga, A., 1991. Corporate bond price discrepancies in the dealer and exchange markets. Journal of Fixed Income 1, 7–16.