The Informational Role of Stock and Bond Volume

Kerry Back
Jones Graduate School of Business and Department of Economics, Rice University

Kevin Crotty
Jones Graduate School of Business, Rice University

In a Kyle (1985) model, the sign of the correlation between a firm's debt and equity returns is the same as the sign of the cross-market Kyle's lambda. The sign is positive (negative) if private information concerns the mean (risk) of the firm's assets. We show empirically that information conveyed by order flows is primarily about asset means. The cross-market lambdas are quite large; consequently, the portions of bond and stock returns explained by order flows are highly correlated, even though the order flows themselves are virtually uncorrelated. (*JEL* G12, G14)

Because they are derivatives with monotone payoffs written on a firm's assets, the prices of corporate debt and equity claims should respond in the same direction to information about the mean value of a firm's assets, but, because the bond payoff is concave and the stock payoff convex, they should respond in opposite directions to information about the risk of a firm's assets. The nature of information—whether it is predominantly about means or risks—is difficult to ascertain in general, but we show that it is possible to determine the nature of private information that arrives to the market via order flows.

In a Kyle (1985) model of informed trading in a firm's debt and equity, different types of private information have different implications for the sign of the cross-market Kyle's lambdas: the cross-market lambdas are positive if information is primarily about the asset mean and negative if information is primarily about the asset risk. We show empirically that the cross-market lambdas are positive, implying that private information is primarily about means. In fact, the cross-market lambdas are quite large. Consequently, the parts of the bond and stock returns that are explained by order flows are

We are grateful to Shmuel Baruch, David De Angelis, Jefferson Duarte, Nishad Kapadia, Pete Kyle (TFF discussant), Barbara Ostdiek, Yuhang Xing, participants at the 2014 Texas Finance Festival, two anonymous referees, and the editors—Leonid Kogan and Pietro Veronesi—for helpful comments. Send correspondence to Kerry Back, Jones Graduate School of Business and Department of Economics, Rice University, Houston, TX 77005; telephone: (713) 348-4168. E-mail: Kerry.E.Back@rice.edu.

doi:10.1093/rfs/hhu094 Advance Access publication December 9, 2014

highly correlated (the median correlation is 80%), even though the order flows themselves are virtually uncorrelated. This is in marked contrast to the correlation between total bond and stock returns. Indeed, the low correlation between total bond and stock returns is widely regarded as a puzzle—see, for example, Collin-Dufresne, Goldstein, and Martin (2001) and Kapadia and Pu (2012).

It is quite intuitive that the sign of the cross-market lambdas should depend on the nature of information. If information is about the asset risk, then stock investors have good news when bond investors have bad news, and vice versa. So, when informed bond investors sell, it is good news for stocks; consequently, cross-market lambdas are negative. On the other hand, by a symmetric argument, when information is about the asset mean, cross-market lambdas are positive. While this seems like a very general argument, it is not trivial to establish within a theoretical Kyle (1985) model. The difficulty is that, because corporate debt and equity are derivatives of a firm's assets, it is not reasonable to assume that their values are joint normally distributed. Indeed, it is the convexity/concavity of payoffs that cause bonds and stocks to respond in opposite directions to information about risks. To incorporate these features of security payoffs, we solve a continuous-time Kyle model with a single informed trader. As shown by Back (1992), such models are tractable even when asset values have non-normal distributions (our innovation is to solve such a model with a discrete rather than continuously distributed asset value).

We document that both cross-market lambdas and own-market lambdas are larger for firms with high-yield debt than for firms with investment-grade debt. In our theoretical model, all lambdas are larger for firms with higher credit spreads when information is primarily about the asset mean but not when information is primarily about the asset risk, so this cross-sectional variation in lambdas reinforces the conclusion that information is primarily about means.

The magnitudes of the cross-market lambdas, relative to own-market lambdas, are nearly identical for high-yield firms and investment-grade firms, but they fell significantly for financial firms during the recent financial crisis. To measure the sign and size of the cross-market lambdas compared with own-market lambdas, we note that, if the order flows were uncorrelated, then the correlation between the parts of the bond and stock returns explained by order flows would be due entirely to the cross-market lambdas. So, we compute the correlation between the explained parts of the stock and bond returns that would be implied by the lambdas if the order flows were uncorrelated and homoscedastic. This equals the cosine of the angle between the rows of the lambda matrix, a metric we call the scaled inner product. This measure of the cross-market lambdas fell dramatically for financial firms during the financial crisis. While there are other possible explanations, this is consistent with there being increased private information about risks for financial firms during the crisis.

To determine whether information about risks is empirically significant, one might be tempted to regress bond returns or spreads on stock return volatility.[1] Our theoretical model shows that it is hazardous to draw inferences from such a regression regarding the importance of information about risks. Specifically, an increase in the perception of future risks can coincide with a reduction in stock return volatility. In our theoretical model, we assume the informed trader has binary information about the mean and/or risk of the future asset value. The market learns through the order flow about which of the two possible information states has been realized. As in Veronesi (1999), standard deviations of price changes are highest when the market is most uncertain about the information state—when both states are considered equally probable. If one of the states is high risk and the other is low risk, then the maximum stock return volatility (standard deviation of log price change) occurs when the probability of the high-risk state is less than $1/2$. For higher probabilities of the high-risk state, information that tends to confirm the high-risk state will lower return volatility. Thus, estimating cross-market lambdas is a more reliable method of determining whether information is about risks or means, though of course this applies only to private information that is manifested in order flows.

Our results hinge on the presence of informed trading in corporate bonds. Evidence that there is a significant amount of informed trading in corporate bonds is provided by Wei and Zhou (2012), Kedia and Zhou (2014), and Han and Zhou (2014). Wei and Zhou (2012) examine corporate bond trading and returns prior to earnings announcements. They find that pre-announcement bond trade imbalance is predictive of earnings surprises and post-announcement bond returns. Kedia and Zhou (2014) obtain similar results for trading and returns of target company bonds prior to acquisition announcements. Han and Zhou (2014) show that measures of adverse selection are significantly related to credit spreads.

Anecdotal evidence of informed trading in corporate bonds includes McCracken (2008), who reports allegations that one or more informed investors shorted Delphi bonds. A particularly interesting case is described by Szockyj (1993). In that case, a number of executives and relatives of the founder of Carl Karcher Enterprises (CKE) were alleged to have traded illegally in both the debt and equity of CKE prior to a negative earnings announcement. Such joint informed trading in debt and equity is particularly likely by investors who hold both securities. Bodnaruk and Rossi (2013) document that institutional investors often simultaneously hold the debt and equity of a firm; these "dual holders" own on average about 10% of the overall shares outstanding for the firms. Acharya and Johnson (2010) also posit that informed trading may occur simultaneously in the debt and equity of a company. They create measures of the ex post likelihood of informed trading in bonds, credit default swaps, stock,

[1] Such regressions are run by Collin-Dufresne, Goldstein, and Martin (2001), Campbell and Taksler (2003), and Cremers et al. (2008), using the VIX, idiosyncratic volatility, and implied volatilities, respectively, as the volatility measure.

and stock options prior to acquisitions by private equity funds and show that they are positively related to the number of potential insiders.

Bond transactions costs are higher for smaller transactions (Schultz, 2001). Our goal is not to explain bid-ask spreads for stocks and bonds but instead to measure the permanent price impacts of bond and stock trades on bond and stock prices. To minimize bid-ask bounce in bond returns, we consider only firms with actively traded bonds, exclude small transactions, and measure returns over hourly or daily time intervals. Han and Zhou (2014) provide evidence that information asymmetry is higher in larger bond transactions, consistent with information-based market microstructure models.

The question of why information is predominantly about asset means is beyond the scope of this paper. It is possible that information about asset risks is inherently more difficult to obtain than information concerning expected values, so markets perceive trading activity as conveying the latter. It is also possible there is some market segmentation and information about risks is exploited in the stock option market rather than in stock and bond markets. We discuss a possible extension of our model to options markets in the conclusion.

1.1 Literature review

While there is a large literature on asymmetric information and capital structure choice, relatively few papers look at informed trading across bond and stock markets. Boot and Thakor (1993) and Fulghieri and Lukin (2001) study competitive rational expectations equilibria, focusing on the effect of security design on the incentives for investors to acquire information. Chang and Yu (2010) analyze a competitive rational expectations model and examine the extent to which managers can learn from informed traders via market prices. Lesmond, O'Connor, and Senbet (2008) analyze a Kyle model but assume riskless debt, so there is informed trading only in the equity market.

Several papers study multi-security Kyle models with the assumption of joint normal asset values, including Caballé and Krishnan (1994), Boulatov, Hendershott, and Livdan (2013), and Pasquariello and Vega (forthcoming). Caballé and Krishnan solve a single-period Kyle model with joint normally distributed asset values. Their main result is that the matrix of Kyle's lambdas is positive definite.[2] Pasquariello and Vega study a two-period version of the Caballé-Krishnan model and show empirically that cross-market lambdas are typically significantly different from zero (frequently negative, but significant). Boulatov, Hendershott, and Livdan modify the two-period model by assuming that market makers can see order flows (and prices) only in their own market. This introduces cross-serial correlation of returns and therefore serial correlation of portfolio returns. They document empirically that institutional order flows lead market returns. Unlike these papers, we relax the assumption

[2] We provide empirical evidence that this is true, and as far as we know, we are the first to show this in any markets.

of joint normality, so that we can analyze informed trading in a firm's debt and equity.

Most bond transactions occur in the over-the-counter market, in which customers contact dealers sequentially. Models of over-the-counter markets based on search theory have been studied by Duffie, Garleanu, and Pedersen (2005), Zhu (2012), and He and Milbradt (2014), among others. Search frictions add costs on top of the purely informational price impacts in Kyle models. We have nothing to say about those additional costs; however, the informational price impacts that we do study seem to be important. For example, we show both theoretically and empirically that stock orders affect bond prices, which would not necessarily be a feature of a pure search model. Thus, the two approaches seem to complement one another.

To our knowledge, this is the first empirical work on net order flows across the corporate stock and bond markets. Previous cross-market studies in this spirit have focused on multiple stocks (Chan, Menkveld, and Yang 2007; Tookes 2008; Andrade, Chang, and Seasholes 2008; Boulatov, Hendershott, and Livdan 2013; Pasquariello and Vega forthcoming), the stock and option markets (Chan, Chung, and Fong 2002; Rourke 2014), or stock and government bond markets (Chordia, Sarkar, and Subrahmanyam 2005). Of course, there are many studies of price impacts of order flows in stock and bond markets separately, including Breen, Hodrick, and Korajczyk (2002), Hasbrouck (2009), and Goyenko, Holden, and Trzcinka (2009) in equity markets and Edwards, Harris, and Piwowar (2007) and Bessembinder, Maxwell, and Venkataraman (2006) in corporate bond markets.

Previous studies of the relationship between debt and equity markets have primarily focused on the degree of comovement and the relative timing of returns in the two markets rather than trading activity. Concerning the extent of comovement, Alexander, Edwards, and Ferri (2000) study cross-market return comovement around events likely to transfer wealth between debt- and equity-holders. They find that debt/equity returns are weakly positively correlated in general but often move in opposite directions around such events. Extending the results of Collin-Dufresne, Goldstein, and Martin (2001), Kapadia and Pu (2012) argue that bonds and stocks move in opposite directions too frequently in the short run and propose limits to arbitrage as an explanation.

The literature contains mixed results regarding the relative informational efficiency of bond and stock markets and whether stock returns lead bond returns. Studies concluding that stock markets lead bond markets include Kwan (1996), Alexander, Edwards, and Ferri (2000), and Downing, Underwood, and Xing (2009). On the other hand, Hotchkiss and Ronen (2002), Ronen and Zhou (2013), and Kedia and Zhou (2014) find that bond markets are as informationally efficient as related equities. In contemporaneous work, Mao (2013) studies the relative contribution of the bond market to price discovery using the information share approach of Hasbrouck (1995) and finds bond markets contribute 13% to price discovery. The relative informational efficiency of equity and credit default swap (CDS) markets has also been studied. Acharya and Johnson (2007) find that there is incremental information in CDS prices relative to equities, but this is disputed by Hilscher, Pollet, and Wilson (forthcoming), who find that the stock market leads the CDS market.

We show that the portions of returns explained by order flows exhibit similar levels of cross-serial correlations. The unexplained portion of stock returns leads the unexplained component of bond returns with a median cross-serial correlation of 10%, but the opposite cross-serial correlation is not generally different from zero. Thus, we conclude that the equity market leads the bond market for the unexplained portions of returns but not for the parts explained by order flows. Unlike previous work, we study cross-market price impacts due to trading activity to discern the nature of private information (means versus risks), which is a possible explanation for the documented low correlation between returns in the stock and bond market.

2. Theory

2.1 Model

The theoretical model is a continuous-time Kyle (1985) model. There is a constant risk-free rate r. A bond and stock of a firm are traded continuously on a time interval $[0,T]$. An announcement affecting the bond and stock values is made after the close of trading at date T. The bond is a zero-coupon bond and matures at T or after. A single risk-neutral trader receives a signal about the announcement at date 0. The signal is binary. We do not assume the informed trader knows the post-announcement values with certainty. His signal can leave residual uncertainty. What is important in the Kyle model are the informed trader's expectations of the post-announcement bond and stock values, conditional on his signal values. We denote the conditional expectations in the two information states by (B_1,S_1) and (B_2,S_2), respectively.

As an illustration, assume the bond matures at the announcement date T, and the announcement is the firm's asset value $\tilde{v}$ at date T. Assume $\tilde{v}$ is drawn from a mixture of lognormals: $\log\tilde{v}=\tilde{\mu}-\tilde{\sigma}^2/2+\tilde{\sigma}\xi$, where ξ is a standard normal and the mean-risk pair $(\tilde{\mu},\tilde{\sigma})$ is either (μ_1,σ_1) or (μ_2,σ_2). In this example, the informed trader's signal is $(\tilde{\mu},\tilde{\sigma})$. Denote the face value of the debt by D, and normalize so that there is a single share of the bond and a single share of stock outstanding. Abstracting from taxes and deadweight bankruptcy costs, the expected post-announcement bond and stock values conditional on the informed trader's signal are

$$B_i=\int_{-\infty}^{\infty}\min\left(\mathrm{e}^{\mu_i-\sigma_i^2/2+\sigma_i x},D\right)\mathrm{n}(x)\mathrm{d}x, \tag{1a}$$

$$S_i=\int_{-\infty}^{\infty}\left(\mathrm{e}^{\mu_i-\sigma_i^2/2+\sigma_i x}-D\right)^+\mathrm{n}(x)\mathrm{d}x, \tag{1b}$$

where n denotes the standard normal density function.

In addition to the informed trades, there are also liquidity trades in the bond and stock, modeled as correlated Brownian motions. Risk-neutral market makers observe net orders in both the bond and the stock and compete to fill them, pushing the market prices to discounted expected values, conditional on the information in the order flows.[3] As in all Kyle models, the informed trader understands how prices depend on the order history and takes it into account when optimizing. Also, again as in all Kyle models, the announcement resolves the information asymmetry, and positions can be liquidated frictionlessly at the post-announcement values.

Let Z^b and Z^s denote the Brownian motions representing the cumulative liquidity trades in the bond and stock respectively. Let

$$\Sigma = \begin{pmatrix} \sigma_b^2 & \rho\sigma_b\sigma_s \\ \rho\sigma_b\sigma_s & \sigma_s^2 \end{pmatrix}$$

denote the instantaneous covariance matrix of $Z=(Z^b, Z^s)$. Let X^b and X^s denote the cumulative trades of the informed trader, and set $Y^i = X^i + Z^i$, for $i \in \{b, s\}$. Market makers observe the vector stochastic process $Y=(Y^b, Y^s)$. Let P^i denote the market price of asset i for $i \in \{b, s\}$. We will derive an equilibrium in which $P_t^i = p^i(t, Y_t)$ for some functions p^i. This means that the prices depend only on cumulative orders at each date and not on the complete prior history of orders. We use the symbols P_T^i to denote the prices at the close of trading, prior to the announcement. There can be jumps at the announcement if the informed trader's signal leaves residual uncertainty about the post-announcement values.

2.2 Information about the asset mean versus the asset risk

Set $\Delta_B = B_2 - B_1$ and $\Delta_S = S_2 - S_1$. The sign of $\Delta_B \Delta_S$ depends on whether private information is primarily about the asset mean or primarily about the asset risk. To see this, consider the lognormal example discussed earlier. The Black-Scholes call option formula implies that the bond and stock expected values B_i and S_i defined by Equation (1) are given by $B_i = f_B(\mu_i, \sigma_i)$ and $S_i = f_S(\mu_i, \sigma_i)$, where

$$f_B(\mu,\sigma) = D\mathrm{N}\left(\frac{\mu - \log D - \sigma^2/2}{\sigma}\right) + \mathrm{e}^{\mu}\mathrm{N}\left(-\frac{\mu - \log D + \sigma^2/2}{\sigma}\right),$$

$$f_S(\mu,\sigma) = \mathrm{e}^{\mu}\mathrm{N}\left(\frac{\mu - \log D + \sigma^2/2}{\sigma}\right) - D\mathrm{N}\left(\frac{\mu - \log D - \sigma^2/2}{\sigma}\right)$$

[3] Our assumption that market makers observe order flow in both markets is not extreme. In the model, market makers can infer orders from price changes, so the assumption that we actually need is post-trade transparency. See Section 3.3 for a discussion of transparency in these markets.

with N denoting the standard normal distribution function. The Black-Scholes formulas for the delta and vega of a call option imply

$$\frac{\partial f_B(\mu,\sigma)}{\partial \mu}=\mathrm{e}^{\mu}\mathrm{N}\left(-\frac{\mu-\log D+\sigma^2/2}{\sigma}\right),\quad \frac{\partial f_S(\mu,\sigma)}{\partial \mu}=\mathrm{e}^{\mu}\mathrm{N}\left(\frac{\mu-\log D+\sigma^2/2}{\sigma}\right),$$

$$\frac{\partial f_B(\mu,\sigma)}{\partial \sigma}=-\mathrm{e}^{\mu}\mathrm{n}\left(\frac{\mu-\log D+\sigma^2/2}{\sigma}\right),\quad \frac{\partial f_S(\mu,\sigma)}{\partial \sigma}=\mathrm{e}^{\mu}\mathrm{n}\left(\frac{\mu-\log D+\sigma^2/2}{\sigma}\right).$$

Thus, the bond and stock values are each increasing in the asset mean e^{μ}. However, the bond value is decreasing and the stock value increasing in the asset risk σ. If, for example, $\sigma_1=\sigma_2$, so the private information is purely about the asset mean e^{μ}, then both B_i and S_i are higher in the high-mean state, implying $\Delta_B\Delta_S>0$. On the other hand, if $\mu_1=\mu_2$, so information is purely about the asset risk, then B_i is higher in the low-risk state and S_i is higher in the high-risk state, so $\Delta_B\Delta_S<0$.

For a general binary signal, if information is primarily about the asset mean, then B and S move together and are higher in the state with the higher mean. However, if information is primarily about the asset risk, then B is higher in the low-risk state, and S is higher in the high-risk state, due to the concavity of the former and convexity of the latter. Therefore, the sign of $\Delta_B\Delta_S$ is positive if information is primarily about the asset mean and negative if information is primarily about the asset risk. As an example, suppose the mean and risk move in opposite directions and state 2 is the low-mean/high-risk state. Then, the bond price will be lower in state 2 than in state 1, so $\Delta_B<0$. The value of Δ_S is ambiguous. If the risk information is more important than the mean information for the stock price, then $\Delta_S>0$. In this case, we say that information is primarily about the asset risk, because $\Delta_B\Delta_S<0$. On the other hand, if the mean information is more important for the stock price, then $\Delta_S<0$, and we say that information is primarily about the asset mean.

A more general definition of information being about the asset mean versus the asset risk that is consistent with ours is the following. Consider a general distribution of the asset value and a general signal distribution for the informed trader, and let $\tilde{B}$ and $\tilde{S}$ denote the expected bond and stock values conditional on the informed trader's signal. Then, we could say that information is primarily about the asset mean if $\mathrm{cov}(\tilde{B},\tilde{S})>0$ and primarily about the asset risk if $\mathrm{cov}(\tilde{B},\tilde{S})<0$. We discuss the sign of the cross-market lambda in a model with a general distribution of the asset value and a general signal distribution in Section 2.10.

2.3 Definition of equilibrium

One requirement for equilibrium is that the bond and stock prices equal discounted expected values, conditional on the market makers' information and given the trading strategy of the informed trader. Let π_t denote the conditional probability of the second information state given the market makers' information at date t, that is, conditional on the history of Y through date t. Prices equal discounted expected values if

$$P_t^b = \mathrm{e}^{-r(T-t)}(B_1 + \Delta_B \pi_t), \tag{2a}$$

$$P_t^s = \mathrm{e}^{-r(T-t)}(S_1 + \Delta_S \pi_t). \tag{2b}$$

The other requirement for equilibrium is that the informed trades are optimal. Let θ_t^b and θ_t^s denote the rate at which the informed trader trades the bond and stock, respectively.[4] These rates have to be adapted to the information possessed by the informed trader, which is the binary signal and the history of Z (in the Appendix, we show that the equilibrium prices reveal enough about the history of Z to enable the informed trader to implement the equilibrium trading strategy, so it is not necessary that he observe Z directly). The informed trader chooses the rates to maximize his conditional expected profit

$$\mathsf{E}\int_0^T \left(B_i - \mathrm{e}^{r(T-t)} P_t^b\right)\theta_t^b \,\mathrm{d}t + \mathsf{E}\int_0^T \left(S_i - \mathrm{e}^{r(T-t)} P_t^s\right)\theta_t^s \,\mathrm{d}t, \tag{3}$$

in each information state $i \in \{1,2\}$. Here, the expectations are over the path of the liquidity trades Z. Of course, the rates of trade θ can and will differ across the two information states. Also, the informed trader takes it as given that the prices in Formula (3) are determined as $P_t^k = p^k(t, Y_t)$ for $k \in \{b, s\}$, with the functions p^k being regarded by the informed trader as exogenous. In the optimization, we also assume that the informed trader is constrained to satisfy the “no doubling strategies” condition introduced in Back (1992), meaning that the strategy must be such that

$$\mathsf{E}\int_0^T \left[p^k(t, Y_t)\right]^2 \mathrm{d}t < \infty$$

for $k \in \{b, s\}$ and in each information state.

2.4 Equilibrium

Let Δ denote the column vector $(\Delta_B, \Delta_S)'$, and define

$$\phi = \sqrt{\Delta' \Sigma \Delta}. \tag{4}$$

To exclude trivial cases, assume $\phi \neq 0$. The parameter ϕ is the instantaneous standard deviation of the stochastic process $\Delta' Y_t$, which plays an important role in the equilibrium—see Equation (6) below. For any real number a, define

$$\kappa(t, a) = \frac{\mathrm{e}^{-r(T-t)}}{\phi\sqrt{T-t}} \times \mathrm{n}\left(\frac{a}{\phi\sqrt{T-t}}\right), \tag{5}$$

where, as before, n denotes the standard normal density function. Let N denote the standard normal distribution function. The proof of the following is in Appendix A.

[4] It is without loss of generality to assume that the informed trades are of order $\mathrm{d}t$ ($\mathrm{d}X = \theta\,\mathrm{d}t$ for some θ), because Back (1992) shows that jumps and Brownian motion components are suboptimal.

Proposition 1. Let π_0 denote the unconditional probability of the second state. Set $\alpha = \phi\sqrt{T}\,\mathrm{N}^{-1}(\pi_0)$. There is an equilibrium in which the prices satisfy Equation (2) with

$$\pi_t = \mathrm{N}\left(\frac{\alpha + \Delta' Y_t}{\phi\sqrt{T-t}}\right). \tag{6}$$

The equilibrium trading strategy is $\theta_t = q_i(t, Y_t)$ in information state i, where

$$q_1(t,y) = \frac{1}{T-t}\mathsf{E}[Z_T - y \mid Z_t = y, \Delta' Z_T < -\alpha], \tag{7a}$$

$$q_2(t,y) = \frac{1}{T-t}\mathsf{E}[Z_T - y \mid Z_t = y, \Delta' Z_T > -\alpha]. \tag{7b}$$

The bond and stock prices evolve as $\mathrm{d}P_t = r P_t \mathrm{d}t + \Lambda_t \mathrm{d}Y_t$, where Λ_t is the singular matrix $\kappa(t, \alpha + \Delta' Y_t)\Delta\Delta'$. In this equilibrium, the aggregate order process Y is a Brownian motion with zero drift and instantaneous covariance matrix Σ given the market makers' information. Moreover, there is convergence to strong-form efficiency in the sense that $\lim_{t\to T} P_t^b = B_i$ and $\lim_{t\to T} P_t^s = S_i$ with probability one in each information state i.

Note that the cross-market lambda is $\kappa(t, \alpha + \Delta' Y_t)\Delta_B \Delta_S$, which is a positive multiple of $\Delta_B \Delta_S$. Thus, the cross-market lambda is positive when information is primarily about the asset mean and negative when information is primarily about the asset risk. This is discussed further at the end of this section. The remainder of the section addresses the interpretation and implications of the proposition.

2.5 The distribution of order flows and learning by the market

In the proof of the proposition, we show that the trading strategy defined by Equation (7) has the property that the informed trades cannot be predicted by market makers; that is, $\mathsf{E}_t[\theta_t] = 0$ for all t when the expectation is conditional on market makers' information. This is a standard feature of Kyle models. It implies that the aggregate order process Y is a martingale given market makers' information.

As the proposition states, the instantaneous covariance matrix of Y is Σ, which is also the instantaneous covariance matrix of Z. This follows from the fact that the instantaneous covariance matrix depends only on the Brownian motion components, and the Brownian motion component of Y is Z. Combining this with the martingale property of Y implies that the distribution of changes in Y over discrete time periods is also the same as the distribution of changes in Z over discrete time periods, because the covariance of two martingales (for example, Y^i and Y^j for $i, j \in \{b, s\}$) is the sum (or integral) of the covariances of the martingale increments, by iterated expectations. In particular, the covariance matrix of the vector Y_T is $\int_0^T (\mathrm{d}Y_t)(\mathrm{d}Y_t)' = T\Sigma$. Thus, the distribution of aggregate order flows is determined by the distribution of liquidity trades and does not depend on whether information is about the asset mean or the asset risk. In fact, it is a general property of Kyle models that the nature of information cannot be identified from order flow data alone. The reason is that the informed trader adapts his trades to the liquidity trades. If he has good news about an asset, he will buy it (for most parameter values—see the discussion in the next section). But, he will buy very little if liquidity traders happen to buy a lot. See Back, Crotty, and Li (2014) and the last paragraph of the next section for more on this topic.

The statement in the proposition that Y is a Brownian motion given the market makers' information follows from it being a martingale and from its instantaneous covariance matrix being a constant matrix. This result is called Levy's theorem.

We also show in the proof of the proposition that the trading strategy (7) implies that $\Delta' Y_T < -\alpha$ with probability one in information state 1 and $\Delta' Y_T > -\alpha$ with probability one in information state 2. In equilibrium, market makers are aware that $\Delta' Y_T < -\alpha$ in state 1 and $\Delta' Y_T > -\alpha$ in state 2. They use this fact, the fact that Y is a Brownian motion, and the current value of $\Delta' Y_t$ at each time t to calculate the conditional probabilities of the two states. This produces Equation (6). It also implies that the market learns the information state over time: the conditional probability of state 2 in Equation (6) satisfies $\pi_t \to 0$ with probability one in state 1 and $\pi_t \to 1$ with probability one in state 2. This produces the convergence to strong-form efficiency stated in the proposition.

2.6 Informed trading

A more explicit formula can be obtained for the equilibrium trading strategy defined by Equation (7) by computing the conditional expectations. For each $i \in \{b, s\}$ and each $t < T$, we have the projection formula

$$Z_T^i - Z_t^i = \beta^i (\Delta' Z_T - \Delta' Z_t) + \varepsilon^i ,$$

where

$$\beta^i = \frac{\text{cov}(Z_T^i - Z_t^i, \Delta' Z_T - \Delta' Z_t)}{\text{var}(\Delta' Z_T - \Delta' Z_t)} = \frac{(T-t)\text{cov}(Z_T^i, \Delta' Z_T)}{(T-t)\text{var}(\Delta' Z_T)} = \frac{\text{cov}(Z_T^i, \Delta' Z_T)}{\text{var}(\Delta' Z_T)}, \tag{8}$$

and where ε^i has zero mean and (due to normality) is independent of $\Delta' Z_T - \Delta' Z_t$. We have

$$\mathsf{E}[\varepsilon^i \mid Z_t^i, \Delta' Z_T > -\alpha] = \mathsf{E}[\varepsilon^i] = 0,$$

because Z has independent increments and because ε^i is independent of $\Delta' Z_T - \Delta' Z_t$. Therefore,

$$\mathsf{E}[Z_T^i - Z_t^i \mid Z_t, \Delta' Z_T > -\alpha] = \beta^i \mathsf{E}[\Delta' Z_T - \Delta' Z_t \mid Z_t, \Delta' Z_T > -\alpha],$$

and similarly for the case $\Delta' Z_T < -\alpha$. Applying a standard formula for the mean of a normal random variable conditional on it being above or below a theshold yields the following formulas in the two information states for the informed trading strategy defined by Equation (7) :

$$q_1(t,y) = -\sqrt{\frac{T-t}{2\pi}} \cdot \frac{\phi e^{-f(t,y)^2/2}}{N(-f(t,y))} \cdot \beta, \tag{9a}$$

$$q_2(t,y) = \sqrt{\frac{T-t}{2\pi}} \cdot \frac{\phi e^{-f(t,y)^2/2}}{N(f(t,y))} \cdot \beta, \tag{9b}$$

where $f(t,y) = (\alpha + \Delta' y)/\sqrt{(T-t)\phi^2}$ and $\beta = (\beta^b \; \beta^s)'$.

For typical parameter values, the informed trader trades the assets in the same direction (either buying both or selling both) when information is primarily about the asset mean and trades them in opposite directions (buying one and selling the other) when information is primarily about the asset risk. To see this, note that Equation (9) implies that the signs of the trades are the signs of the β's in information state 2 and the opposite in information state 1. Also, Equation (8) implies that β^b has the sign of $\sigma_b \Delta_B + \rho \sigma_s \Delta_S$ and β^s has the sign of $\sigma_s \Delta_S + \rho \sigma_b \Delta_B$. If the correlation ρ between the liquidity trades is not too large in absolute value, then β^b has the same sign as Δ_B, and β^s has the same sign as Δ_S. In this case, the informed trader always buys undervalued securities and sells overvalued securities. When information is primarily about the asset mean, he buys both securities in the high-mean state and sells both in the low-mean state. On the other hand, when information is primarily about the asset risk, then he buys one security and sells the other.

However, there are parameter values for which the informed trader will buy an overvalued security or sell an undervalued security. Suppose, for example, that information is primarily about the asset mean and state 2 is the high-mean state, so $\Delta_B > 0$ and $\Delta_S > 0$. Suppose that the correlation ρ between the liquidity trades is negative and sufficiently large in absolute value that $\sigma_b \Delta_B + \rho \sigma_s \Delta_S < 0$. Observe that this implies that $\sigma_b \Delta_B < \sigma_s \Delta_S$, so we must have $\sigma_s \Delta_S + \rho \sigma_b \Delta_B > 0$. Therefore, the informed trader buys the stock and sells the bond in state 2, even though both securities are undervalued. Likewise, he sells the stock and buys the bond in state 1, even though both securities are overvalued. Note that by trading the securities in opposite directions in this example, the informed trader causes the correlation of aggregate order flows to be negative, consistent with the negative correlation of liquidity trades. This helps to explain why the aggregate orders have the same distribution as the liquidity trades, as discussed in the previous section. Nevertheless, the informed trades in this example have the effect of eventually increasing both prices in state 2 and decreasing both prices in state 1, with the market eventually learning which state has occurred. This dependence of informed trades on the correlation of liquidity trades is not unique to our model. It also occurs in the single-period model of Caballé and Krishnan (1994).

2.7 Credit spreads and lambdas

The lambda matrix in the proposition depends only on time and on the conditional probability π_t. Moreover, the labeling of the states is irrelevant: the matrix is the same when the probability of the second information state is x as it is when the probability of the first information state is x, for any $0<x<1$. This follows from the definition (5) of $\kappa(t,a)$, Equation (6) for π_t, and the formula $\Lambda_t = \kappa(t,\alpha+\Delta' Y_t)\Delta\Delta'$. Specifically,

$$\begin{aligned}\Lambda_t &= \frac{\mathrm{e}^{-r(T-t)}}{\phi\sqrt{T-t}} \times \mathrm{n}\left(\frac{\alpha+\Delta' Y_t}{\phi\sqrt{T-t}}\right)\Delta\Delta' \\ &= \frac{\mathrm{e}^{-r(T-t)}}{\phi\sqrt{T-t}} \times \mathrm{n}\left(\mathrm{N}^{-1}(\pi_t)\right)\Delta\Delta' \\ &= \frac{\mathrm{e}^{-r(T-t)}}{\phi\sqrt{T-t}} \times \mathrm{n}\left(\mathrm{N}^{-1}(1-\pi_t)\right)\Delta\Delta',\end{aligned}$$

the last equality following from the symmetry of the normal distribution.

In our empirical work, we estimate the impact of order flows on returns instead of price changes, due to the usual stationarity concern. We estimate the matrix Γ, where $\Gamma^{ij} = \Lambda^{ij}/P^i$ for $i, j \in \{b,s\}$. The symmetry of Λ with respect to the probabilities implies an asymmetry of Γ. Suppose that information is primarily about the asset mean, and assume state 2 is the high-mean state, so $B_1 < B_2$ and $S_1 < S_2$. Then, the elements of the Λ matrix are the same when the conditional probability of the first state is 80% as when the conditional probability of the second state is 80%, and the prices are lower in the first case; therefore, the elements of the Γ matrix are larger when the conditional probability of the first state is 80% than when the conditional probability of the second state is 80%. Thus, they are larger when the bond price is lower—equivalently, when the credit spread is higher.

The above result is a comparative statics result, comparing Γ_t for two different values of π_t. It has cross-sectional implications. Consider two firms with the same values of $\Delta_B > 0$ and $\Delta_S > 0$ and the same face value of debt. At any date t, the firm with the higher value of π_t will have both the lower credit spread and the smaller Γ matrix. Thus, if information is primarily about the asset mean ($\Delta_B \Delta_S > 0$), then our model predicts that the coefficients in regressions of returns on order flows will be larger for firms with lower credit ratings. On the other hand, if information is primarily about the asset risk, then the low-bond-value state is the high-stock-value state, so the elements of the Γ matrix corresponding to the stock return should be smaller for lower rated firms. In our empirical work, we find that all of the elements of Γ are larger for lower rated firms, consistent with private information being primarily about asset means.

2.8 Return volatilities

Set $A_t = \alpha + \Delta' Y_t$ and

$$dW_t = \frac{1}{\phi} dA_t .$$

Combining Equations (2) and (6), we see that all investors can infer the value of A from the equilibrium prices. Hence, W is observable to the market at large. The proposition shows that Y and hence W is a continuous martingale given the market's information. Because $(dW)^2 = dt$, it follows that W is a standard Brownian motion given the market's information. The formulas $dP_t = r P_t dt + \Lambda_t dY_t$ and $\Lambda_t = \kappa(t, \alpha + \Delta' Y_t)\Delta\Delta'$ imply that

$$dP_t^b = r P_t^b dt + \phi \Delta_B \kappa(t, A_t) dW_t , \tag{10a}$$

$$dP_t^s = r P_t^s dt + \phi \Delta_S \kappa(t, A_t) dW_t . \tag{10b}$$

The function $\kappa(t, \cdot)$ is proportional to the standard normal density function. Therefore, the standard deviations of the absolute price changes are strictly decreasing in $|A_t|$. Combining this fact with Equation (6), we see that the standard deviations decrease as the probability π_t of the second state decreases toward zero or increases toward one and are highest when $\pi_t = 1/2$. Thus, the standard deviations are increasing in the degree of uncertainty about the information state rather than being increasing in the probability of the high-risk state, as one might have conjectured. This phenomenon also appears in Veronesi (1999).

The return volatilities (standard deviations of log price changes) are

$$\frac{\phi|\Delta_B|\kappa(t, A_t)}{P_t^b}, \qquad \frac{\phi|\Delta_S|\kappa(t, A_t)}{P_t^s},$$

respectively. Suppose information is about the asset risk and state 1 is the high-risk state, so $B_1 < B_2$ and $S_1 > S_2$. Then, the maximum bond return volatility will occur for $\pi_t < 1/2$ and the maximum stock return volatility will occur for $\pi_t > 1/2$. Thus, paradoxically, the maximum stock return volatility occurs when the low-risk state is more probable. This highlights the danger in using changes in stock return volatility to proxy for information about the asset risk.

2.9 Return correlation and cross-market lambda

Our binary information model is too simple to be a good model of the stock-bond correlation. However, it suffices to illustrate our main point: both the correlation and the cross-market lambda are (i) positive if information is primarily about the asset mean, (ii) negative if information is primarily about the asset risk, and (iii) zero or near zero if information is mixed. In fact, it is clear from Equation (10) that the bond-stock correlation is +1 if $\Delta_B \Delta_S > 0$, -1 if $\Delta_B \Delta_S < 0$, or 0 if $\Delta_B \Delta_S = 0$. The first two cases correspond to information that is primarily about the asset mean or primarily about the asset risk, respectively, as we have already discussed. As an example of the third case, suppose state 1 is low mean and high risk and state 2 is high mean and low risk. Then, $B_2 > B_1$, because both the high mean and the low risk are good for the bond value. However, if the parameters are fixed precisely right, then we can obtain $S_1 = S_2$. In this case, the stock is locally risk-free, hence uncorrelated with the bond. Also, if state 1 is low mean and low risk and state 2 is high mean and high risk, then $S_2 > S_1$, but we could have $B_1 = B_2$, in which case the bond is locally risk-free.

In more general models, the matrix Λ_t should be nonsingular; hence, the bond and stock returns should be less than perfectly correlated. In a previous version of this paper, we assumed continuous information (with the informed trader knowing the asset value $\tilde{v}$ exactly) and proved that Λ_t is in fact positive definite for each t. This is consistent with the results of Back (1993) and Caballé and Krishnan (1994).

2.10 Cross-market lambdas and the nature of information for general signals

Consider a Kyle model in which the asset value has an arbitrary distribution and there is a single informed trader who has an arbitrarily distributed signal about the values of a stock and bond. Let $(\tilde{B}, \tilde{S})$ denote the expectations of the bond and stock values given the informed trader's information. Assume for simplicity that the risk-free rate is zero and the liquidity trades are uncorrelated across markets, so Σ is a diagonal matrix. Assume the following three conditions, which seem to be robust properties of Kyle models:

- (i) The price processes P^b and P^s are martingales—this follows from risk-neutral market makers.
- (ii) The terminal prices P^b_T and P^s_T satisfy $P^b_T = \tilde{B}$ and $P^s_T = \tilde{S}$ with probability one—this follows from the informed trader not "leaving money on the table"; see Back (1992).
- (iii) The price changes satisfy $\mathrm{d}P = \Lambda \,\mathrm{d}Y$ for a symmetric matrix Λ—symmetry of the lambda matrix holds in our model and is also established in multi-asset Kyle models by Back (1993) and Caballé and Krishnan (1994).

From standard properties of martingales, the covariance between P^b_T and P^s_T equals the expected cumulative covariation:

$$\mathrm{cov}(P^b_T, P^s_T) = \mathsf{E}\int_0^T (\mathrm{d}P^b_t)(\mathrm{d}P^s_t) = \mathsf{E}\int_0^T \psi_t \,\mathrm{d}t,$$

where ψ_t is the off-diagonal element of $\Lambda_t \Sigma \Lambda_t$. Therefore,

$$\mathrm{cov}(\tilde{B}, \tilde{S}) = \mathsf{E}\int_0^T \lambda^{bs}_t (\sigma^2_s \lambda^{ss}_t + \sigma^2_b \lambda^{bb}_t)\,\mathrm{d}t.$$

The right-hand side of this is a weighted average of the cross-market lambda, weighted by the positive process $\sigma^2_s \lambda^{ss} + \sigma^2_b \lambda^{bb}$. Thus, the cross-market lambda is on average positive when $\text{cov}(\tilde{B}, \tilde{S}) > 0$—which is true when information is primarily about the asset mean—and is on average negative when $\text{cov}(\tilde{B}, \tilde{S}) < 0$—which is true when information is primarily about the asset risk. If information is mixed, we could have $\text{cov}(\tilde{B}, \tilde{S}) = 0$, in which case the average cross-market lambda is zero. Thus, our result about the nature of information and the sign of the cross-market lambda does not depend on our binary signal assumption, provided only that conditions (i)–(iii) hold.

3. Data

3.1 Sample construction

In order to understand the information content of stock and bond order flow, we seek to empirically characterize the price impact matrix. The empirical analysis utilizes nine years of transactions data from the stock and corporate bond markets. Our sample period begins in July 2002, which corresponds to the initiation of bond transaction reporting requirements for broker-dealers by the National Association of Securities Dealers (NASD), and runs through June 2011. As of July 2007, the reporting requirements were imposed by the Financial Industry Regulatory Authority (FINRA). Stock market trade and quote data are obtained from NYSE TAQ, and the corporate bond transactions data are from FINRA TRACE.[5] Bond characteristics are obtained from the Mergent Fixed Income Security Database (FISD). Daily equity returns and shares outstanding are obtained from CRSP. The data appendix provides the details on data matching and filters. We restrict the universe of firms to those for which there is an active market for both equity and debt instruments. Specifically, to be included in the sample, we require that a firm has an observed bond return on at least 1,000 of the 2,221 possible trading days. The resulting sample contains observations on 221 firms.

Summary statistics concerning trading activity in our sample are given in Table 1. The median stock market capitalization is $14 billion, so the sample is skewed toward large firms. This is not surprising given the requirement that firms have actively traded public debt securities. The median par amount of debt outstanding is about $5 billion. Despite the fact that we consider relatively liquid bonds, firms in our sample still have fewer days with a bond trade than a stock trade. Moreover, on days with trades, the number of bond trades is dwarfed by the number of equity trades. The median daily number of bond trades is 19, while the median daily number of stock trades is over 12,000. However, bond trades are for substantially larger dollar amounts, so dollar volume across the two markets is much more similar than is the number of trades. The median firm

[5] A previous version of this paper used data from 2008 through 2011. FINRA publicly disclosed trade direction indicators starting in November 2008. In this version, we employ the Enhanced TRACE database, which includes trade direction indicators back to 2002. The Enhanced TRACE also does not top-censor volume for investment-grade (high-yield) trades of $5 million ($1 million).

Table 1
Debt and equity trading activity statistics

	Bond trading activity				
	Volume ($ thousands)	Principal ($ millions)	# Daily trades	# Trade days	Turnover (0.001s)
Mean	39,897	17,376	48	1,989	5.0
SD	81,027	45,129	116	307	3.7
P25	12,227	2,434	10	1,851	2.6
P50	19,833	5,309	19	2,151	4.0
P75	34,710	9,633	43	2,209	6.4
N	221	221	221	221	221
	Stock trading activity				
	Volume ($ thousands)	Market cap ($ millions)	# Daily trades	# Trade days	Turnover (0.001s)
Mean	185,539	28,784	18,509	2,117	10.3
SD	224,389	41,735	19,618	254	7.6
P25	57,245	5,596	7,420	2,201	5.6
P50	107,148	14,071	12,623	2,220	8.1
P75	226,131	31,851	21,677	2,221	12.2
N	221	221	221	221	221

The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Stock and bond volume are in thousands of dollars. Debt outstanding and market capitalization are in millions of dollars. Bond and stock turnover is multiplied by 1,000.

has almost $20 million of bonds trade on an average day compared with stock dollar volume of over $100 million. Median daily bond turnover is around half that for equity.

We calculate bond returns as $\frac{P_t - P_{t-1}}{P_{t-1}}$, where P_t is the last recorded transaction price on day t. Trading costs are much higher for smaller volume trades in the corporate bond market (Schultz, 2001, Edwards, Harris, and Piwowar, 2007). To mitigate possible bid-ask bounce effects on the return series, we use prices from transactions of at least $100,000 when creating the return series for each bond. Most firms in our sample have multiple bond issues outstanding. To avoid overweighting firms with many bonds outstanding, we aggregate bond trading activity by firm. The unit of observation is thus a firm-day. For returns, we calculate a weighted average bond return using daily returns from any bonds with observed returns on a given day. The weights are calculated as the principal amount outstanding of the observed bond as a fraction of the total principal outstanding for all bonds with observed activity on day t. We keep firm-day observations for which the weighted-average bond return is defined; that is, where at least one bond of the firm traded on day $t-1$ and the same bond traded on day t.

3.2 Measurement of signed volume

To estimate signed volume, we must sign the aggressive side of trades. For equity trades, this is done using the Lee and Ready (1991) algorithm. That is, trades above (below) the prevailing quote midpoint are considered buys (sells). If a trade occurs at the midpoint, then the trade is classified as a buy (sell) if the trade price is greater (less) than the previous transaction price. We define stock order imbalance as shares bought less shares sold divided by total shares outstanding.

For bond trades, FINRA collects information on the reporting party's side of the trade or an indication that the trade is an interdealer trade. Since reporting firms are FINRA-member broker-dealers, trades are classified as buys (sells) if the TRACE trade indicator is "S" ("B"). We identify trades as belonging to a round-trip transaction if two or more trades are observed at the same volume but different prices within fifteen minutes of each other.[6] For trades that are matched as part of a round-trip trade, the first trade is presumed to be the aggressive trade. Signed volume for a firm's debt is defined as the difference in par values bought and sold (across all outstanding bonds for the firm), divided by the total face value outstanding. For both signed volume measures, we normalize by shares or principal outstanding to allow for cross-sectional comparison.

A brief discussion on measurement error of the order imbalance measures is warranted. Stock and bond signed volume are both measured with error. The TAQ data omits odd-lot trades—trades for less than 100 shares. O'Hara, Yao, and Ye (2014) document that an increasing number of trades are odd-lot trades. The equity order imbalance measure we use here is the imbalance in the number of shares traded, rather than the imbalance in the number of trades. Since odd-lot trades concern small numbers of shares by definition, the share order imbalance measure is less prone to error than the number of trades order imbalance. O'Hara, Yao, and Ye (2014) find that only 3.33% of order imbalances are misclassified when using volume measures. Trade misclassification due to the Lee and Ready (1991) algorithm could also lead to measurement error in stock order imbalance.

For bonds, we expect less measurement error given the availability of order direction indicators. However, signed order flow could be incorrectly measured for round-trip trades if we do not correctly identify matching trades or if the broker-dealer reports the trade in the wrong sequence. We believe measurement error is not first order given that we find the bond within-market price impact is positive as expected.

Summary statistics for variables used in our analysis are found in Table 2. To reduce the effect of outliers, all variables are winsorized at the 1%/99% levels. As expected, stock returns are more volatile than bond returns. On the other hand, bond signed order flow exhibits more variation than does stock signed order flow.

6 Round-trip trades have been used to estimate transactions costs by Bessembinder, Maxwell, and Venkataraman (2006), Goldstein, Hotchkiss, and Sirri (2007), and Feldhütter (2012) for corporate bonds and Green, Hollifield, and Schürhoff (2007) for municipal bonds.

Table 2
Summary statistics of returns and signed volume

	r_t^b	r_t^s	x_t^b	x_t^s
Mean	3	5	−0.153	0.375
SD	75	233	2.470	1.470
Min	−267	−783	−10.565	−3.924
P25	−26	−100	−0.767	−0.222
P50	1	2	−0.017	0.169
P75	31	109	0.577	0.732
Max	284	823	8.865	7.180
N	334,460	334,460	334,460	334,460

The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume (x_t^i) is normalized by shares (principal) outstanding for the stock (bond) market (both are multiplied by 1,000). Returns (r_t^i) are measured in basis points. All variables are winsorized at the 1%/99% levels.

3.3 Observability of order flow

While market makers in the model observe order flow in both markets, this may not be true in practice. In the model, market makers can infer orders from price changes, so the assumption that we actually need is post-trade transparency. Given the increased use of odd-lot orders that are not captured by the consolidated tape (O'Hara, Yao, and Ye 2014), equity markets have become somewhat less transparent post-trade, but remain fairly transparent. While corporate bond markets have been opaque historically, post-trade bond prices are now transparent due to the TRACE reporting requirement (Edwards, Harris, and Piwowar 2007; Bessembinder, Maxwell, and Venkataraman 2006; Goldstein, Hotchkiss, and Sirri 2007). Corporate bond trade reporting requirements for NASD member dealers' over-the-counter bond transactions were initiated on July 1, 2002. Reported trade information includes bond identifier, date and time of execution, trade size and price, yield, and a buy/sell indicator for the dealer's side of the trade. Not all of the reported information is publicly disseminated. Trade size is censored at \$5 million for investment-grade and \$1 million for high-yield bonds. The buy/sell indicator information was not publicly disseminated until November 3, 2008. The requirement for public dissemination of bond transactions expanded in stages. Initially, only trades in investment-grade bonds with original issuance size of over \$1 billion and trades in a group of 50 high-yield bonds (former FIPS bonds) were publicly reported. On March 3, 2003, TRACE expanded public reporting to cover all bonds rated A or higher with issuance of at least \$100 million. On April 14, 2003, 120 representative BBB bonds became publicly transparent. Finally, on October 1, 2004, practically all over-the-counter corporate bond trades were publicly reported, with full implementation by February 7, 2005. The required timeliness of reporting also increased over this time period. Initially, dealers were required to report trades within 75 minutes of the trade. This declined to 45 minutes on October 1, 2003, and again to 30 minutes on October 1, 2004. Beginning July 1, 2005, dealers have had to report trades within fifteen minutes. Our primary empirical results are based on daily returns and signed volume, so they are unlikely to be affected by reporting lags for corporate bonds.

Boulatov, Hendershott, and Livdan (2013) motivate the use of a daily vector autoregression through a Kyle model in which market makers cannot immediately observe order flow in correlated assets. The unobservability induces order flows to exhibit positive cross-serial correlation with subsequent returns. Given the relatively short reporting lags for bonds and stocks during our sample, our assumption that order flows are observable within the daily frequency seems reasonable. Nonetheless, we investigate lead-lag relationships between returns and order flows in a modified vector autoregression framework in Section 4.4. Empirically, the cross-asset price impacts are primarily due to contemporaneous order flows rather than a lead-lag relationship, suggesting that observability is not an issue in our setting.

3.4 Additional data

Section 2.7 discusses how the estimated price impact matrix could change with credit quality. In our empirical analysis, we use credit ratings as our measure of credit quality. We collect credit ratings from S&P and Moody's from FISD. The bond ratings are translated into a numerical value ranging from 21 for AAA/Aaa down to 1 for C or below. The firm-level rating from each agency is calculated as the average bond rating weighted by the amount outstanding for each bond. A firm is classified as investment grade if the average of its resulting Moody's and S&P ratings is greater than or equal to 12. If the average rating is less than 12, the firm is classified as high yield. This corresponds to the usual investment-grade/high-yield boundary of BBB-/Baa3. Approximately 70% of our firm-day observations are classified as investment grade.

We also estimate price impacts using hourly returns and order imbalances. Specifically, we examine firms with bonds that trade in at least five hourly bins on average over the sample period. We construct hourly bond and stock returns and order imbalances between the hours of 10 a.m. and 4 p.m. We require that firms have at least 1,000 hourly observations to be included in the analysis. For intraday bins without bond trading, we assume that the bond price is unchanged since the last transaction price; that is, the bond return is zero for that hour. The sample does not include overnight returns. Naturally, a smaller number of firms satisfy these requirements; our hourly sample includes 77 firms.

4. Empirical Analysis

4.1 Empirical specification and theoretical predictions

Our primary empirical model is a linear relation between returns and order flows in the two markets:

$$r_{it} = \Gamma x_{it} + \epsilon_{it} \quad \Leftrightarrow \quad \begin{array}{ccccccc} r_{it}^b & = & \Gamma^{bb} x_{it}^b & + & \Gamma^{bs} x_{it}^s & + & \epsilon_{it}^b \\ r_{it}^s & = & \Gamma^{sb} x_{it}^b & + & \Gamma^{ss} x_{it}^s & + & \epsilon_{it}^s \end{array} \tag{11}$$

where x^s (x^b) is daily signed volume normalized by shares (principal) outstanding for the stock (bond) of firm i, and r_{it}^j is the return of security j for firm i. We use returns rather than price changes as the dependent variable because the sizes of price changes are sensitive to the absolute price level and to facilitate cross-sectional comparison.[7] Note that Γ^{ij} is an estimate of Λ^{ij}/P^i from the model.[8] In the following subsections, we implement the linear specification (11), and we also include lagged returns and order flows as explanatory variables in a modified VAR.

Several aspects of the Γ and Λ matrices are of interest. The magnitudes of the diagonal elements reflect the liquidity in each market. That topic has been extensively studied before. We are interested in the off-diagonal elements. According to the theory, the off-diagonal elements are positive if information is primarily about the mean of the firm's asset value and negative if information is primarily about the risk of the firm's asset value. In addition to the signs of the off-diagonal elements, we are interested in their magnitudes relative to the magnitudes of the diagonal elements. A simple measure would be $\Lambda^{bs}/\sqrt{\Lambda^{bb}\Lambda^{ss}}$ or $\Lambda^{sb}/\sqrt{\Lambda^{bb}\Lambda^{ss}}$, which are equal when Λ is symmetric, as established in the proposition. Instead, for reasons explained below, we use the following as a measure of the sign and magnitude of the off-diagonal elements:

$$\frac{\Lambda^{bb}\Lambda^{sb}+\Lambda^{bs}\Lambda^{ss}}{\sqrt{((\Lambda^{bb})^2+(\Lambda^{bs})^2)((\Lambda^{sb})^2+(\Lambda^{ss})^2)}}. \tag{12}$$

Assuming Λ is nonnegative definite and symmetric, we have

$$-1 \le \frac{\Lambda^{bs}}{\sqrt{\Lambda^{bb}\Lambda^{ss}}} = \frac{\Lambda^{sb}}{\sqrt{\Lambda^{bb}\Lambda^{ss}}} \le 1, \tag{13}$$

and the measure (12) takes its minimum value -1 when Λ^{bs} is negative and large in the sense that the ratio in condition (13) equals -1; it is 0 when $\Lambda^{bs}=0$; and it takes its maximum value +1 when Λ^{bs} is positive and large in the sense that the ratio in Condition (13) equals +1.

The measure (12) is the inner product of the rows of Λ divided by the product of the norms of the rows. We call it the scaled inner product. It equals the cosine of the angle between the rows of Λ. More concretely, it is the correlation that the price changes (the elements of Λx_{it}) would have if the bond and stock signed volumes x_{it}^b and x_{it}^s were uncorrelated and homoscedastic.

[7] We obtain similar results using excess returns rather than raw returns, suggesting that the relationship we study is capturing the firm-specific information that is the focus of our study. Excess returns are calculated using an equal-weighted index of returns in each market.

[8] In the model, price impacts are stochastic; that is, Λ is time varying and depends on cumulative order flow. We estimate the empirical equivalent of $\mathrm{d}P_t = \Lambda(t, Y_t)\mathrm{d}Y_t$ where $\mathrm{d}P$ is sampled at the daily or hourly frequency and normalized by price. Our estimates represent an average price impact across firms and time. The (untabulated) results are similar if we include quadratic signed order-flow terms in the return regressions to account for the fact that we sample at a lower frequency than continuous time.

Note that the measure (12) is unchanged if we calculate it using the Γ matrix instead of the Λ matrix, because the divisions by prices in going from Λ to Γ cancel in the numerator and denominator of the ratio (12). This invariance with respect to the substitution of Γ for Λ is the reason we use the ratio (12) rather than the ratio in (13) as a measure of the sign and magnitude of the cross-market lambdas. The invariance implies that we can calculate the measure by estimating the specification (11). We test whether it is equal to zero. We also test whether it was smaller during the financial crisis, whether it is smaller for financial firms than for nonfinancials, and whether it is different prior to earnings announcements.

We are also interested in whether the price impact matrix Λ is positive definite.[9] Positive definiteness of Λ has two implications. First, it means that price impact costs are positive. In a single security model with the change in price being proportional to the trade size ($\Delta P = \lambda \Delta Y$), the price impact cost is $(\Delta P)(\Delta Y) = \lambda(\Delta Y)^2 > 0$. In a multiasset model, positivity of the price impact cost of any vector of trades ΔY is equivalent to positive definiteness of the price impact matrix Λ: $(\Delta P)'(\Delta Y) = (\Delta Y)'\Lambda(\Delta Y) > 0$. Second, positive definiteness of Λ means that each market is relatively more sensitive to its own orders. Consider bond and stock orders each scaled to have a unit impact on the stock price, meaning $\Delta Y^b = 1/\Lambda^{sb}$ and $\Delta Y^s = 1/\Lambda^{ss}$. If Λ is positive definite, the bond market will move more from the bond order than from the stock order, meaning

$$\frac{\Lambda^{bb}}{\Lambda^{sb}} > \frac{\Lambda^{bs}}{\Lambda^{ss}}.$$

Likewise, for bond and stock orders scaled to have a unit impact on the bond price, the stock market will move more from the stock order than from the bond order.

Recall that Γ is related to the price impact matrix Λ by $\Lambda^{ij} = P^i \Gamma^{ij}$. Assuming Λ is symmetric, it is positive definite if and only if $\Lambda^{bb} > 0$ and $\Lambda^{bb}\Lambda^{ss} > \Lambda^{bs}\Lambda^{sb}$, which is equivalent to $\Gamma^{bb} > 0$ and $\Gamma^{bb}\Gamma^{ss} > \Gamma^{bs}\Gamma^{sb}$. To test whether these conditions hold, we test the null hypothesis

$$\Gamma^{bb} = 0 \quad \text{or} \quad \Gamma^{bb}\Gamma^{ss} - \Gamma^{bs}\Gamma^{sb} = 0$$

formulated as

$$\Gamma^{bb}\left(\Gamma^{bb}\Gamma^{ss} - \Gamma^{bs}\Gamma^{sb}\right) = 0. \tag{14}$$

A rejection of this hypothesis also implies a rejection of the null that the scaled inner product of the rows of Γ is equal to -1 or 1. We also test the positive definiteness of Λ without assuming symmetry; see footnote 12.

9 As discussed in Section 2, the matrix Λ is singular when there is binary information, but this is special to the case of binary information. In a previous version of this paper with continuous information, we verified the theoretical positive definiteness of Λ for trading a company's debt and equity. Other research on multisecurity Kyle models also establishes positive definiteness (Back, 1993, Caballé and Krishnan, 1994).

In addition to tests related to the Γ matrix, we are interested in time-series properties of the returns r_{it}, the explained parts of the returns $\hat{\Gamma}x_{it}$, and the residuals $\hat{\epsilon}_{it}$. In particular, we are interested in the correlations of the different parts of the bond and stock returns and in the cross-serial correlations. These shed light on the extent to which the markets are integrated, where price discovery occurs, and the extent to which information is about means versus risks.

4.2 Seemingly unrelated regressions

Table 3 presents seemingly unrelated regression (SUR) estimates of specification (11), which is the multiasset equivalent of the price impact specification used by Breen, Hodrick, and Korajczyk (2002). In all regressions, standard errors are adjusted for within-firm serial correlation and daily across-firm correlation.[10]

We report estimates for the full sample and for the subsamples of investment-grade firms and high-yield firms. In each case, all elements of the estimated Γ matrix are positive and highly significant. The positive signs of the off-diagonal elements show that the order flows predominantly convey information about asset means rather than information about asset risks. The estimates for the subsample of high-yield firms are all at least twice the size of the estimates for the investment-grade firms, except for the within-market equity price impact. These results are also consistent with information being primarily about asset means, as discussed in Section 2.7. All price impact estimates are positive in firm-by-firm time-series regressions (Panel D of Table 3) as well.

To get a sense of the economic magnitude of the estimates, note that our signed volume measures are scaled such that a unit change is equivalent to an order imbalance of 0.1% of shares outstanding (principal outstanding) for stock (bond) order flow. The standard deviation of equity signed volume is 1.47 in our sample, meaning 0.147% of shares outstanding. A stock order imbalance of this size is associated with an equity price impact of 41 bps (1.47×28.2) and a bond price impact of 5 bps (1.47×3.5). For the median firm in our sample, this order imbalance represents 670,000 shares. The existing microstructure literature usually estimates equity price impacts at a higher frequency (often five-minute intervals). Our daily within-equity-market price impact is of similar magnitude to the five-minute price impact of a 1,000-share order, estimates of which have ranged from roughly 10 to 30 bps.[11] Not surprisingly, an order imbalance of 1,000 shares has almost no price impact at the daily frequency.

---

[10] Specifically, the variance-covariance matrix is estimated as $V_{\text{Firm\&Time}} = V_{\text{Firm}} + V_{\text{Time}} - V_{\text{Robust}}$, as proposed by Cameron, Gelbach, and Miller (2011) and Thompson (2011) and discussed in Petersen (2009). As a practical matter, within-firm serial correlation has the largest effect on standard errors, which are largely unchanged when additionally clustered by time.

[11] See, for example, Breen, Hodrick, and Korajczyk (2002), Hasbrouck (2009), and Goyenko, Holden, and Trzcinka (2009).

Table 3
Estimation of Γ

	r_t^b		r_t^s
Panel A. Full sample			
x_t^b	1.61***		1.30***
	[13.76]		[4.40]
x_t^s	3.48***		28.16***
	[10.66]		[17.07]
Observations	334,460		334,460
R^2	0.008		0.034
Scaled inner product		0.93***	
Panel B. Investment-grade firms			
x_t^b	1.05***		0.76**
	[9.35]		[2.29]
x_t^s	1.95***		25.82***
	[5.27]		[9.40]
Observations	228,284		228,284
R^2	0.002		0.019
Scaled inner product		0.90***	
Panel C. High-yield firms			
x_t^b	2.24***		1.83***
	[13.86]		[4.25]
x_t^s	4.25***		30.03***
	[9.33]		[15.77]
Observations	102,445		102,445
R^2	0.016		0.051
Scaled inner product		0.91***	
Panel D. Firm-by-firm regressions			
x_t^b	1.72***		1.65***
	[15.95]		[5.96]
x_t^s	2.15***		35.52***
	[5.52]		[19.56]
Observations	334,460		334,460
R^2	0.011		0.043
Scaled inner product		0.48***	

This table presents estimates of seemingly unrelated regressions of bond and stock returns on signed volume in the stock and bond markets. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. A firm-day is classified as investment grade if the average of its S&P and Moody's bond ratings is greater than or equal to BBB-/Baa3 on a numeric scale. Firm-days with an average rating less than BBB-/Baa3 are classified as high yield. For Panels A–C, standard errors are calculated clustering by firm and day. Panel D reports average coefficients from firm-by-firm time-series regressions. t-statistics are in brackets (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

On the other hand, a daily bond order imbalance of 2.47 (its standard deviation) is associated with a bond price impact of 4 bps (2.47 × 1.61) and an equity price impact of 3 bps (2.47 × 1.30). For the median firm in our sample, this corresponds to an order imbalance of approximately $13 million. Edwards, Harris, and Piwowar (2007) and Bessembinder, Maxwell, and Venkataraman (2006) allow bond trade execution costs to vary with trade size. Their estimates of the half-spread for a $10 million trade range from 4 to 18 bps, and our daily price impact estimate falls within this range.

The scaled inner product of the rows of the estimated Γ matrix is 0.93. We can strongly reject the null hypothesis that the scaled inner product is zero for the full sample and both credit-quality subsamples. The scaled inner products are similar for investment-grade (0.90) and high-yield firms (0.91). Furthermore, for the full sample and for each credit-quality subsample, the Wald test strongly rejects the null hypothesis (14)—$p < 0.001$ in all cases—so we conclude that Λ is positive definite.[12] This also implies a rejection of the null that the scaled inner product is equal to 1. Though the scaled inner product is less than 1, so the fitted bond and stock returns are not perfectly correlated, the large point estimate implies that the correlation between the fitted bond and stock returns should be large. This is indeed the case, as we demonstrate next.

Table 4 reports cross-sectional distributions of the time-series correlations of the components of returns explained by signed volume along with the correlations of returns, signed order flows, and the residual components of returns (Panel A). The correlations of daily stock and bond returns are moderate, with a median value of 11%. However, the fitted values from specification (11) are highly correlated. The median correlation is 80%, and the 5th percentile correlation is 62%. These high correlations are consistent with private information being primarily about means, as explained in Section 2.9. On the other hand, the residual components of returns are not highly correlated. The residual components could reflect public information. If so, the low correlation implies that public information is mixed, being partly about risks and partly about means. The low correlation of the residuals could also reflect market segmentation, as suggested by Collin-Dufresne, Goldstein, and Martin (2001) and Kapadia and Pu (2012) based on analyses of total stock and bond returns.

We also present cross-serial correlations in Panel B of Table 4. We find evidence that stock returns lead bond returns, but little evidence of the opposite relationship. We account for this dynamic relationship in the vector autoregression analysis in Section 4.4. Interestingly, this lead-lag relationship is confined to the portions of returns unexplained by signed volume. The cross-serial correlations of the explained portions of returns are remarkably symmetric when alternating the lagged market assignment, indicating that price discovery is not dominated by either market with respect to information conveyed by order flows.

12 We have also verified the positive definiteness of Λ without assuming symmetry. In the absence of symmetry, Λ is positive definite if and only if $\Lambda + \Lambda'$ is symmetric and positive definite. This is equivalent to $\Gamma^{bb} > 0$ and $4\Gamma^{bb}\Gamma^{ss} > \left(\sqrt{\frac{P_b}{P_s}}\Gamma^{bs} + \sqrt{\frac{P_s}{P_b}}\Gamma^{sb}\right)^2$. In our sample, the ratio $\frac{P_b}{P_s}$ has a median value of about 3. We can also reject the null

$$\Gamma^{bb}\left[4\Gamma^{bb}\Gamma^{ss} - \left(3^{0.5}\Gamma^{bs} + 3^{-0.5}\Gamma^{sb}\right)^2\right] = 0.$$

Table 4
Correlation statistics

Percentile:	5th	25th	Median	75th	95th
Panel A. Contemporaneous correlations					
$\rho(r_t^b, r_t^s)$	−0.114	−0.036	0.111	0.261	0.541
$\rho(x_t^b, x_t^s)$	−0.060	−0.028	−0.006	0.016	0.042
$\rho(\hat{r}_t^b, \hat{r}_t^s)$	0.615	0.717	0.796	0.874	0.973
$\rho(\hat{\epsilon}_t^b, \hat{\epsilon}_t^s)$	−0.115	−0.030	0.107	0.249	0.539
Panel B. Cross-serial correlations					
$\rho(r_t^b, r_{t-1}^s)$	−0.010	0.049	0.102	0.159	0.214
$\rho(r_{t-1}^b, r_t^s)$	−0.069	−0.017	0.007	0.036	0.089
$\rho(\hat{r}_t^b, \hat{r}_{t-1}^s)$	0.065	0.127	0.187	0.251	0.400
$\rho(\hat{r}_{t-1}^b, \hat{r}_t^s)$	0.073	0.124	0.184	0.253	0.389
$\rho(\hat{\epsilon}_t^b, \hat{\epsilon}_{t-1}^s)$	−0.012	0.049	0.101	0.154	0.213
$\rho(\hat{\epsilon}_{t-1}^b, \hat{\epsilon}_t^s)$	−0.066	−0.018	0.010	0.039	0.093

This table presents cross-sectional distributions of correlations of various time series across markets and time periods. Panel A shows correlations of stock and bond returns (r_t^i), signed volume (x_t^i), and the explained and residual components of returns. Panel B shows cross-serial correlations of returns and the decomposition of returns based on order flows. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume (x_t^i) is normalized by shares (principal) outstanding for the stock (bond) market (both are multiplied by 1,000). The explained portion of returns ($\hat{r}_t^i$) is measured as the fitted value from specification (11) presented in Table 3. The residual component ($\hat{\epsilon}_t^i$) is computed from the same model. Returns and signed volume are winsorized at the 1%/99% levels.

Table 5 reports the estimate of specification (11) from the panel of hourly returns and order flows for our sample of 77 firms. For the full sample, all estimated price impacts are positive, and all are statistically significant except for the response of stock returns to bond signed volume, which has a t-statistic of 1.3. The within-market price impacts are much larger than the cross-market impacts, the within-stock-market price impact being the largest. However, we can still reject the null that the scaled inner product of the rows of the Γ matrix is zero, although the estimated scaled inner product is much smaller (0.14). The credit-quality subsample analysis indicates that the positive cross-market effects are strongest in the high-yield firms. As at the daily frequency, we reject the null hypothesis (14) for the full sample and both credit quality subsamples ($p < 0.001$), so we can conclude that Λ is positive definite.

4.3 Subsample analysis

The previous section establishes that markets interpret order flow as conveying information that primarily concerns asset means. In this section, we ask whether this result holds for various subsamples in which it is plausible that investors may possess different types of private information. A natural place to start is determining how much time-series variation exists for our price impact estimates, particularly whether there are subsamples with negative cross-market price impacts.

Table 5
Hourly estimation of Γ

	r_t^b		r_t^s
Panel A. Full sample			
x_t^b	2.91***		0.36
	[9.17]		[1.30]
x_t^s	0.40***		38.50***
	[2.78]		[8.20]
Observations	587,334		587,334
R^2	0.003		0.021
Scaled inner product		0.14***	
Panel B. Investment-grade firms			
x_t^b	2.46***		0.21
	[7.83]		[0.73]
x_t^s	0.18		33.21***
	[1.06]		[5.81]
Observations	512,881		512,881
R^2	0.002		0.013
Scaled inner product		0.08	
Panel C. High-yield firms			
x_t^b	4.88***		1.09
	[14.78]		[1.51]
x_t^s	0.82***		48.44***
	[3.72]		[6.19]
Observations	73,949		73,949
R^2	0.007		0.051
Scaled inner product		0.19***	
Panel D. Firm-by-firm regressions			
x_t^b	5.34***		1.13
	[7.52]		[1.38]
x_t^s	−0.04		35.38***
	[−0.20]		[6.42]
Observations	587,334		587,334
R^2	0.006		0.033
Scaled inner product		−0.01	

This table presents estimates of seemingly unrelated regressions of bond and stock returns on signed volume in the stock and bond markets. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 77 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. A firm-day is classified as investment grade if the average of its S&P and Moody's bond ratings is greater than or equal to BBB–/Baa3 on a numeric scale. Firm-days with an average rating less than BBB–/Baa3 are classified as high yield. For Panels A–C, standard errors are calculated clustering by firm and day. Panel D reports average coefficients from firm-by-firm time-series regressions. t-statistics are in brackets (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

Figure 1 shows the time series of estimates of specification (11) over rolling two-year windows. We find no evidence of negative cross-market price impacts over our sample. However, the estimates do exhibit time-series variation. In particular, both own-market and cross-market lambdas increased during the financial crisis. To analyze this further, we estimate the specification (11) separately during the pre-crisis, crisis, and post-crisis periods. Because financial firms were at the heart of the crisis, we also estimate the specification (11)

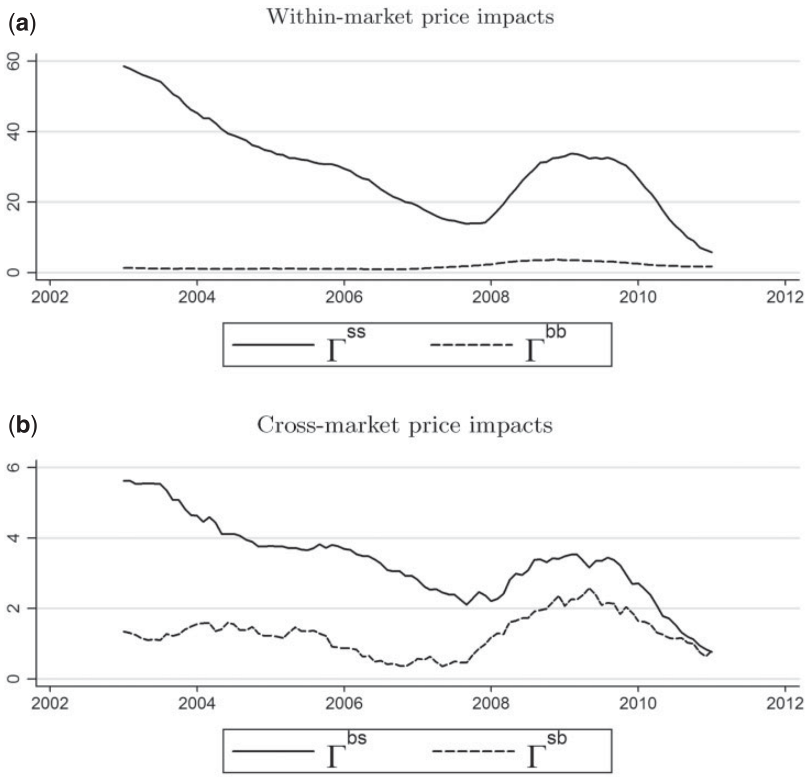

Figure 1
Time series of price impact coefficients
This figure plots the time series of price impact estimates from seemingly unrelated regressions of bond and stock returns on signed volume in the stock and bond markets. The monthly time series is estimated using 24-month overlapping windows. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels.

separately in each period for financial and nonfinancial firms.[13] We take the crisis time period to be July 2007 to June 2009. These dates correspond to the collapse of the Bear Stearns subprime funds and the NBER-dated end of the recession, respectively.

Table 6 reports the scaled inner product for each subsample. There was a significant drop in this measure of the cross-market lambdas for financial firms during the crisis. It also dropped for nonfinancials from the pre-crisis period to the crisis and again from the crisis to the post-crisis period, though the latter decline is not statistically significant. The drop during the crisis for financial firms is consistent with there being more information about risks in the order

[13] We classify financial firms as those with SIC codes between 6000 and 6999.

Table 6
Scaled inner product by industry and time-period

	Full Sample	Non-financial	Financial	$\Delta_{Industry}$
Pre-crisis	0.98***	0.98***	0.87***	0.11
	[0.00]	[0.00]	[0.00]	[0.30]
Crisis	0.73***	0.79***	0.37**	0.42**
	[0.00]	[0.00]	[0.04]	[0.02]
Post-crisis	0.73***	0.69***	0.88***	−0.18
	[0.00]	[0.00]	[0.00]	[0.22]
$\Delta_{Crisis-Pre}$	−0.25***	−0.19***	−0.50**	
	[0.00]	[0.00]	[0.01]	
$\Delta_{Post-Pre}$	−0.24***	−0.29***	0.01	
	[0.00]	[0.00]	[0.97]	
$\Delta_{Post-Crisis}$	0.00	−0.10	0.51**	
	[0.97]	[0.34]	[0.04]	

This table presents estimates of the scaled inner product of the rows of the estimated Γ matrix from seemingly unrelated regressions of bond and stock returns on signed volume in the stock and bond markets. Signed volume is interacted with indicator variables for financial firms and the crisis period. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Financial firms are those with SIC codes between 6000 and 6999. The crisis period includes observations from July 2007 to June 2009. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. p-values for a test that an estimate is equal to 0 are presented in brackets and are calculated using a covariance matrix allowing for clustering by firm and day (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

flows for financial firms during the crisis. Another possible explanation is that the bond and stock markets were less integrated during this time. There is evidence that the corporate bond and credit default swap (CDS) markets became less integrated during the crisis. Bai and Collin-Dufresne (2013) report that, on average, bond spreads exceeded CDS spreads by hundreds of basis points during the crisis, and they attribute that gap to a limited supply of arbitrage capital for purchasing corporate bonds. It is not clear why a limited supply of capital for buying the physical to arbitrage a physical/derivative spread would cause disintegration of the market for two physicals (stocks and bonds), so it seems plausible that the change in the cross-market bond/stock lambda could stem from a different cause. Our model suggests that the cause could be an increase in private information about risks. In any case, the scaled inner product is significantly positive for each time and industry subsample, indicating that the predominant type of information concerned asset means.

The informational environment of a firm is potentially different prior to an earnings announcement (Korajczyk, Lucas, and McDonald 1991). To see if there are differences in the responses of returns to order flows in such periods, we estimate the specification (11) separately for days within a two-week period prior to an announcement and all other days (Table 7).[14] All price impacts are positive in both periods, although the effect of bond order flows on equity returns is not statistically significant prior to earnings announcements.

[14] We thank an anonymous referee for suggesting this analysis.

Table 7
Estimation of Γ around earnings announcements

	r_t^b		r_t^s
Panel A. Non-pre-earnings announcement days			
x_t^b	1.57***		1.40***
	[12.41]		[4.40]
x_t^s	3.39***		27.92***
	[10.82]		[16.55]
Observations	284,003		284,003
R^2	0.007		0.034
Scaled inner product		0.93***	
Panel B. Pre-earnings announcement days			
x_t^b	1.84***		0.74
	[11.43]		[1.38]
x_t^s	4.05***		29.59***
	[7.50]		[15.03]
Observations	50,457		50,457
R^2	0.010		0.034
Scaled inner product		0.92***	
Panel C. Pre-earnings — Same ex post price changes			
x_t^b	1.81***		0.47
	[9.49]		[0.78]
x_t^s	4.56***		28.84***
	[8.22]		[13.70]
Observations	33,919		33,919
R^2	0.012		0.035
Scaled inner product		0.94***	
Panel D. Pre-earnings — Opposite ex post price changes			
x_t^b	1.92***		1.52
	[6.45]		[1.47]
x_t^s	2.70***		31.60***
	[3.60]		[10.79]
Observations	16,538		16,538
R^2	0.006		0.033
Scaled inner product		0.84***	

This table presents estimates of seemingly unrelated regressions of bond and stock returns on signed volume in the stock and bond markets. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. The estimation is done separately for non-pre-earnings intervals (Panel A) and the two-week interval preceding earnings announcements (Panel B). The pre-earnings estimation is also conditioned on whether the ex post price reactions move in the same or opposite directions. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. Standard errors are calculated clustering by firm and day. t-statistics are in brackets (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

Moreover, the scaled inner products are similar for pre-announcement days (0.92) and all other days (0.93).

The information content of earnings announcements can differ from instance to instance; some announcements may contain information that primarily concerns asset risks, while others concern asset means. In practice, market makers should treat pre-earnings order flow as concerning the predominant form of information. In our sample, approximately 30% of the quarterly earnings announcements are associated with opposite-direction three-day cumulative bond and stock returns from $t-1$ to $t+1$, consistent with information about means being the predominant type; however, some announcements apparently do contain information about risks rather than means. The scaled inner product is 0.84 (0.94) for pre-announcement periods with subsequent divergent (concordant) price reactions, which is consistent with there being more private information about risks prior to announcements with divergent reactions, but the difference between the scaled inner products is not statistically significant ($p=0.18$).

4.4 Vector autoregressions

The cross-serial correlation analysis in Section 4.2 shows evidence of a possible dynamic relationship between returns and order flows. In this subsection, we check that our results are robust to such dynamics by allowing returns and order flows to depend on their lagged values. We employ the extended Hasbrouck (1991) vector autoregression (VAR) framework used by Chan, Chung, and Fong (2002) and Chan, Menkveld, and Yang (2007). It differs from a traditional VAR in that signed order flows are presumed to affect contemporaneous returns.

Table 8 reports estimates of an extended VAR for returns and signed volume. Let $\mathbf{r_{i,t}}=[r^b_{i,t} \quad r^s_{i,t}]'$ represent returns on the bond and stock over interval t and $\mathbf{x_{i,t}}=[x^b_{i,t} \quad x^s_{i,t}]'$ represent bond and stock signed volume over interval t. We estimate the system:

$$\begin{aligned}\mathbf{r_{i,t}} &= \mathbf{a_1 r_{i,t-1}}+\ldots+\mathbf{a_p r_{i,t-p}}+\Gamma\mathbf{x_{i,t}}+\mathbf{b_1 x_{i,t-1}}+\ldots+\mathbf{b_p x_{i,t-p}}+\epsilon_{\mathbf{1,i,t}} \\ \mathbf{x_{i,t}} &= \mathbf{c_1 r_{i,t-1}}+\ldots+\mathbf{c_p r_{i,t-p}}+\mathbf{d_1 x_{i,t-1}}+\ldots+\mathbf{d_p x_{i,t-p}}+\epsilon_{\mathbf{2,i,t}}\end{aligned} \tag{15}$$

where $\mathbf{a_1},...,\mathbf{a_p},\Gamma,\mathbf{b_1},...,\mathbf{b_p},\mathbf{c_1},...,\mathbf{c_p},\mathbf{d_1},...\mathbf{d_p}$ are 2×2 matrices of coefficients. Γ contains the coefficients of primary interest. We estimate the model with $p=5$ lags.[15] Table 8 shows that the coefficients of signed bond and stock volume in the return equations are all positive and highly significant. We continue to reject the null that the scaled inner product of the rows of the estimated Γ matrix is equal to 0 at the 0.1% level. We can also strongly reject the null hypothesis (14)—$p<0.001$—so we can again conclude that Λ is positive definite.

We calculate impulse response functions as in Hasbrouck (1991), who argues for persistent price impacts as evidence of informed trading. Figure 2 shows the cumulative price responses due to a one-standard-deviation shock to bond or stock order flows. The impulse response functions are constant after about five trading days, which indicates permanent price effects. A one-standard-deviation shock to bond order flow leads to similar price impacts in both the stock and bond market (Panel 2a) of around 5–6 basis points. A similar shock to stock order flow produces a permanent stock return of over 50 bps and a permanent

15 Results are qualitatively similar for different lag lengths.

Table 8
Relationship between returns and signed volumes of corporate bonds and stocks

	Dependent variables			
	r_t^b	r_t^s	x_t^b	x_t^s
x_t^b	1.47***	1.39***		
	[11.08]	[3.84]		
x_t^s	3.50***	36.04***		
	[9.52]	[14.47]		
r_{t-1}^b	−0.08***	0.08***	−0.32***	0.07
	[−7.46]	[2.80]	[−2.91]	[1.12]
r_{t-2}^b	0.02**	0.04	−0.42***	0.14**
	[2.02]	[1.43]	[−4.13]	[2.17]
r_{t-1}^s	0.04***	−0.02	0.14***	−0.10***
	[15.73]	[−1.60]	[5.14]	[−3.75]
r_{t-2}^s	0.01***	−0.02	0.05*	−0.10***
	[6.72]	[−1.11]	[1.81]	[−3.75]
x_{t-1}^b	0.08	−0.26	45.14***	−3.70**
	[0.87]	[−1.01]	[10.39]	[−2.49]
x_{t-2}^b	0.23***	0.35	21.29***	−5.30***
	[2.59]	[1.34]	[5.45]	[−3.67]
x_{t-1}^s	0.27	−7.23***	−12.91**	238.52***
	[1.38]	[−8.00]	[−2.36]	[37.15]
x_{t-2}^s	−0.48**	−3.91***	−16.11***	123.85***
	[−2.39]	[−5.22]	[−3.06]	[25.30]
Observations	211,997	211,997	211,997	211,997
R^2	0.034	0.041	0.006	0.228
Scaled inner product	0.94***			

This table presents estimates of a modified VAR containing returns and signed volume from both the stock and bond markets. Contemporaneous signed volume is presumed to affect returns as in Hasbrouck (1991) and Chan, Chung, and Fong (2002). The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. Standard errors are clustered by firm and by day. Coefficients in the order flow equations (x^b and x^s) have been multiplied by 1,000. Five lags of each variable are included in the regressions (estimates for lags 3–5 are omitted for space). t-statistics are in brackets (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

bond return of around 10 bps (Panel 2b). In sum, the permanent within- and cross-market effects are consistent with both bond and stock order flows having information content and with the information being predominantly about asset means rather than asset risks.

As discussed in the introduction, a number of studies have addressed the relative informational efficiency of the two markets with mixed results. The coefficients on lagged returns in our VAR speak to this debate (Table 8). In our sample, both lagged cross-market returns are positively associated with returns in each market. This is inconsistent with a single market dominating price discovery. Another interesting result concerns the dynamics of signed order flow within and across markets. Signed order flow in each market exhibits

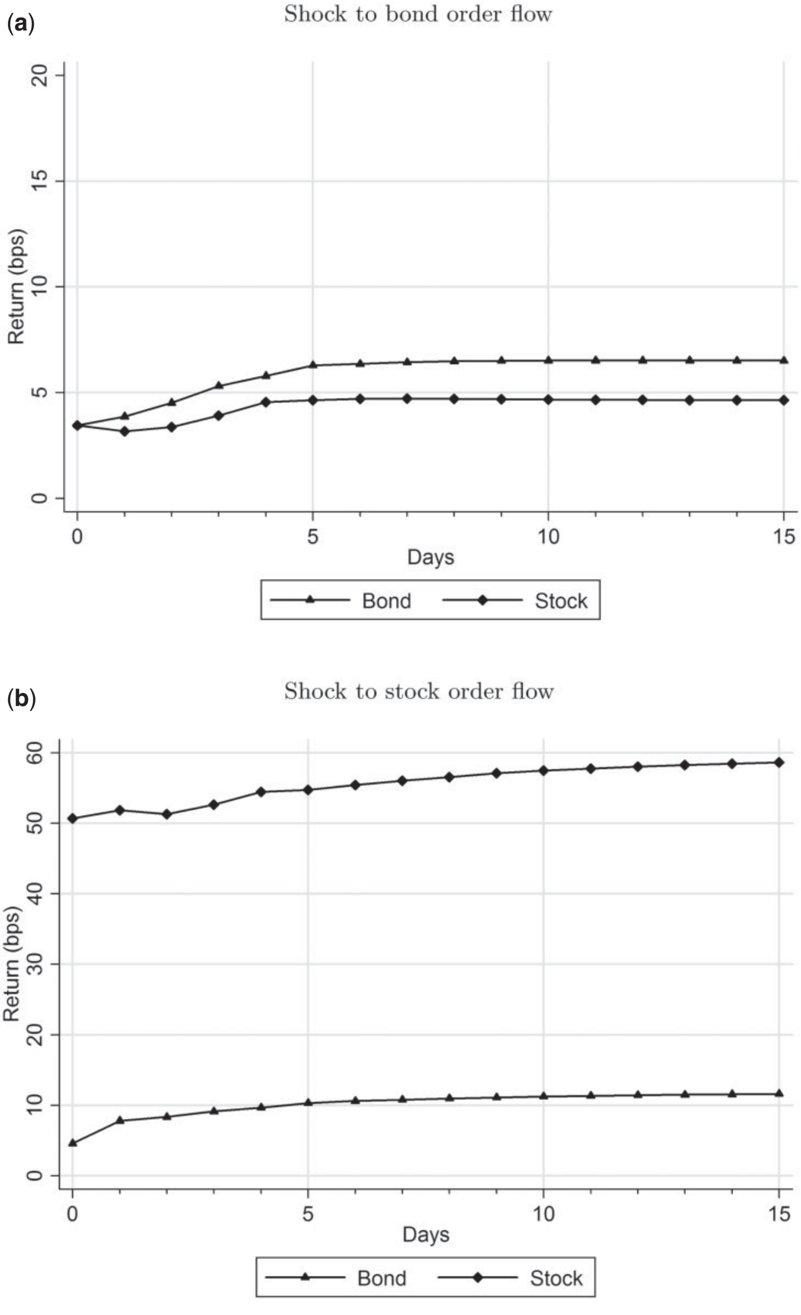

Figure 2
Impulse response functions
This figure plots impulse response functions for bond and stock returns based on the daily VAR model. The functions plot the cumulative return response to a one-standard-deviation shock to signed volume in each market. The impulse response functions use parameter estimates reported in Table 8 using the extended Hasbrouck (1991) VAR model.

Table 9
High-yield bonds versus investment-grade bonds

	Dependent variables			
	r_t^b	r_t^s	x_t^b	x_t^s
x_t^b	0.95***	0.87**		
	[6.48]	[2.05]		
x_t^b * HY Dummy	1.38***	1.63**		
	[4.71]	[2.17]		
x_t^s	1.95***	32.74***		
	[5.00]	[8.39]		
x_t^s * HY Dummy	2.69***	7.27*		
	[4.64]	[1.70]		
r_{t-1}^b	−0.07***	0.08***	−0.38***	0.06
	[−6.34]	[2.91]	[−3.10]	[0.90]
r_{t-2}^b	0.02*	0.04	−0.45***	0.16**
	[1.95]	[1.34]	[−3.79]	[2.24]
r_{t-1}^s	0.04***	−0.02	0.13***	−0.11***
	[15.18]	[−1.63]	[4.93]	[−4.17]
r_{t-2}^s	0.01***	−0.02	0.05**	−0.11***
	[6.09]	[−1.21]	[1.96]	[−3.79]
x_{t-1}^b	0.13	−0.24	43.03***	−3.85**
	[1.32]	[−0.84]	[9.03]	[−2.54]
x_{t-2}^b	0.20**	0.15	19.60***	−5.78***
	[2.17]	[0.51]	[4.26]	[−3.67]
x_{t-1}^s	0.41**	−6.92***	−12.18**	237.83***
	[2.14]	[−6.90]	[−2.24]	[33.15]
x_{t-2}^s	−0.52**	−4.44***	−12.04**	120.07***
	[−2.55]	[−5.68]	[−2.20]	[23.65]
Observations	181,626	181,626	181,626	181,626
R^2	0.035	0.042	0.006	0.220
Scaled inner product – IG	0.92***			
Scaled inner product – HY	0.93***			
Δ_{HY-IG}	0.01			

This table presents estimates of a modified VAR containing returns and signed volume from both the stock and bond markets. Contemporaneous signed volume is presumed to affect returns as in Hasbrouck (1991) and Chan, Chung, and Fong (2002), and is interacted with an indicator for high-yield debt. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. A firm-day is classified as investment grade if the average of its S&P and Moody's bond ratings is greater than or equal to BBB–/Baa3 on a numeric scale. Firm-days with an average rating less than BBB–/Baa3 are classified as high yield. Standard errors are clustered by firm and by day. Coefficients in the order flow equations (x^b and x^s) have been multiplied by 1,000. Five lags of each variable are included in the regressions (estimates for lags 3–5 are omitted for space). t-statistics are in brackets (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

positive serial correlation. In addition, the cross-market serial correlations of order flows are negative.

Table 9 presents estimates of the VAR including interactions of signed bond and stock volumes with a high-yield dummy variable. For the high-rated firms, all elements of the estimated Γ matrix are positive and significant. Moreover, the interaction terms are all positive and significant. This corroborates our previous

Table 10
Daily order imbalance—number of trades

	Dependent variables			
	r_t^b	r_t^s	$x_t^{0,b}$	$x_t^{0,s}$
$x_t^{0,b}$	3.82***	7.90***		
	[5.19]	[3.14]		
$x_t^{0,s}$	26.89***	385.00***		
	[4.52]	[4.98]		
r_{t-1}^b	−0.07***	0.08***	−0.13***	0.01
	[−7.35]	[2.97]	[−4.74]	[1.41]
r_{t-2}^b	0.02**	0.04	−0.10***	0.02***
	[1.97]	[1.45]	[−4.27]	[3.33]
r_{t-1}^s	0.04***	−0.02	0.02	−0.03***
	[15.72]	[−1.26]	[1.57]	[−8.26]
r_{t-2}^s	0.02***	−0.02	0.01	−0.01***
	[6.80]	[−1.07]	[1.30]	[−2.69]
$x_{t-1}^{0,b}$	−0.42	−0.98	197.58***	0.89*
	[−1.05]	[−0.91]	[9.14]	[1.90]
$x_{t-2}^{0,b}$	1.03***	−0.20	125.10***	0.82
	[3.43]	[−0.13]	[7.20]	[1.21]
$x_{t-1}^{0,s}$	0.56	−86.11***	53.45	261.48***
	[0.29]	[−3.62]	[1.56]	[18.52]
$x_{t-2}^{0,s}$	−7.84***	−40.88**	−7.84	105.73***
	[−2.79]	[−2.05]	[−0.43]	[5.23]
Observations	210,110	210,110	210,110	210,110
R^2	0.031	0.041	0.141	0.205
Scaled inner product	0.99***			

This table presents estimates of a modified VAR containing returns and the signed volume from both the stock and bond markets. Contemporaneous signed volume is presumed to affect returns as in Hasbrouck (1991) and Chan, Chung, and Fong (2002). x^0 is the order imbalance between the number of buys less number of sells. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. Standard errors are clustered by firm and by day. Coefficients in the order flow equations ($x^{0,b}$ and $x^{0,s}$) have been multiplied by 1,000. Five lags of each variable are included in the regressions (estimates for lags 3–5 are omitted for space). t-statistics are in brackets (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

result that price impacts of order flows, both within and across markets, are larger for low-rated firms.

Our results are robust to the specification of order imbalance as well. Table 10 reports estimates of the extended VAR with the bond and stock order imbalances measured as the difference between the number of buyer-initiated trades and the number of seller-initiated trades, normalized by the average number of trades in the previous calendar month. This is motivated by Jones, Kaul, and Lipson (1994), who suggest that trade size contains no more information than the number of transactions. Hasbrouck (1991) also uses only the direction of the trade in his VAR analysis, arguing that the model leads to less variable estimates of price impacts. Bond and stock returns are again positively related to both within- and cross-market order flows. The bond trade imbalance variable has

Table 11
Daily order imbalance—direction

	Dependent variables			
	r_t^b	r_t^s	$x_t^{0,b}$	$x_t^{0,s}$
$x_t^{0,b}$	4.19***	4.39***		
	[14.62]	[6.27]		
$x_t^{0,s}$	2.23***	33.29***		
	[6.83]	[12.96]		
r_{t-1}^b	−0.07***	0.08***	−0.05	0.06
	[−7.36]	[2.83]	[−1.29]	[1.46]
r_{t-2}^b	0.02**	0.04	−0.13***	0.08**
	[2.12]	[1.54]	[−3.67]	[2.13]
r_{t-1}^s	0.04***	−0.02*	0.06***	−0.02
	[15.91]	[−1.72]	[5.11]	[−1.00]
r_{t-2}^s	0.01***	−0.02	0.03***	−0.02
	[6.59]	[−1.30]	[2.92]	[−1.59]
$x_{t-1}^{0,b}$	−0.54***	−0.36	29.30***	−4.82**
	[−2.78]	[−0.57]	[12.43]	[−2.19]
$x_{t-2}^{0,b}$	0.36**	0.55	19.31***	−6.39***
	[1.96]	[0.87]	[9.10]	[−2.72]
$x_{t-1}^{0,s}$	0.32*	−5.04***	−1.79	160.57***
	[1.67]	[−6.19]	[−0.71]	[43.63]
$x_{t-2}^{0,s}$	−0.24	−2.44***	−1.54	102.88***
	[−1.18]	[−3.09]	[−0.60]	[28.17]
Observations	211,997	211,997	211,997	211,997
R^2	0.032	0.020	0.003	0.110
Scaled inner product	0.58***			

This table presents estimates of a modified VAR containing returns and the signed volume from both the stock and bond markets. Contemporaneous signed volume is presumed to affect returns as in Hasbrouck (1991) and Chan, Chung, and Fong (2002). x^0 is an indicator variable for the direction of the daily signed order imbalance with $x^0 = 1$ (-1) for a buy (sell) order imbalance. The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 221 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. Standard errors are clustered by firm and by day. Coefficients in the order flow equations ($x^{0,b}$ and $x^{0,s}$) have been multiplied by 1,000. Five lags of each variable are included in the regressions (estimates for lags 3–5 are omitted for space). t-statistics are in brackets (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

a standard deviation of 0.67, so a one-standard-deviation shock to bond trade order imbalance is associated with a 3 (5) bps bond (stock) return. The stock trade imbalance variable has a standard deviation of 0.14, so a one-standard-deviation shock to stock trade order imbalance is associated with a 4 (54) bps bond (stock) return.

Table 11 reports estimates of the extended VAR with the bond and stock order imbalances replaced by their signs. The empirical estimates are again consistent with all of the lambdas being positive and with the matrix of lambdas being positive definite. A day with a buying order imbalance for stocks (bonds) is associated with own-market returns of 33 (4) bps and cross-market returns of 2 (4) bps. As in the other specifications, the impact of bond orders is similar

Table 12
Hourly relationship between returns and signed volumes of corporate bonds and stocks

	Dependent variables			
	r_t^b	r_t^s	x_t^b	x_t^s
x_t^b	2.92***	0.39		
	[9.15]	[1.20]		
x_t^s	0.25*	41.49***		
	[1.76]	[7.73]		
r_{t-1}^b	−0.15***	0.01	0.06	−0.02
	[−30.61]	[1.24]	[1.27]	[−0.68]
r_{t-2}^b	−0.06***	−0.00	−0.02	0.02
	[−22.28]	[−0.16]	[−0.72]	[0.84]
r_{t-1}^s	0.00***	−0.01*	0.02***	0.04***
	[2.59]	[−1.83]	[3.06]	[3.37]
r_{t-2}^s	0.00***	0.00	0.01*	−0.00
	[2.96]	[0.22]	[1.85]	[−0.20]
x_{t-1}^b	−0.49***	0.04	−13.32***	−2.63**
	[−5.26]	[0.15]	[−4.94]	[−2.56]
x_{t-2}^b	−0.36***	0.36	2.32	−2.78**
	[−4.74]	[1.31]	[0.79]	[−2.10]
x_{t-1}^s	0.05	−3.54***	−4.64*	201.26***
	[0.33]	[−2.71]	[−1.76]	[24.52]
x_{t-2}^s	0.10	−4.30***	−2.73	90.36***
	[0.82]	[−4.71]	[−1.20]	[16.11]
Observations	584,724	584,724	584,724	584,724
R^2	0.027	0.022	0.000	0.096
Scaled inner product	0.09*			

This table presents estimates of a modified VAR containing hourly returns and signed volume from both the stock and bond markets. Contemporaneous signed volume is presumed to affect returns as in Hasbrouck (1991) and Chan, Chung, and Fong (2002). The sample runs from July 2002 to June 2011 (2,221 trading days) and contains data from 77 firms. Bond market data is aggregated at the firm level, so the unit of observation is a firm-day. Signed volume is normalized by shares (principal) outstanding for the stock (bond) market. Returns are measured in basis points. All variables are winsorized at the 1%/99% levels. Standard errors are clustered by firm and by day. Coefficients in the order flow equations (x^b and x^s) have been multiplied by 1,000. Five lags of each variable are included in the regressions (estimates for lags 3–5 are omitted for space). t-statistics are in brackets (* $p<0.10$, ** $p<0.05$, *** $p<0.01$).

for both bond and stock returns, while the effect of stock order flows is larger for stock returns than for bond returns.

Table 12 reports estimates of the extended VAR using our panel of hourly returns and signed order imbalances for 77 firms. The price impacts are all positive and significant, with the exception of the effect of bond order flows on stock returns, which is positive but insignificant (t-statistic of 1.2). We can reject the null that the scaled inner product of the estimated Γ matrix is zero with a p-value of 0.050.

5. Conclusion

It has long been understood that the low correlation between corporate bond and stock returns could be due to the nature of information flowing into the market—namely, whether it is about the mean or the risk of the firm's asset value. We show that these two types of information can be distinguished empirically, to the extent that information is private information that arrives to the market via order flows. They can be distinguished because they have different implications for the sign of the cross-market Kyle's lambdas. If private information is primarily about the expected value of the firm's assets, then the equilibrium cross-market price impacts are positive. On the other hand, if private information primarily concerns the variance of the asset value, then cross-market price impacts are negative. If information is mixed, then the cross-market price impacts—and the correlation of bond and stock returns—can be zero or near zero.

We empirically analyze price responses to order flows using transactions data from 2002 through 2011. Returns are positively related to daily order imbalances both within the same market and across markets. These results hold controlling for lagged returns and lagged order imbalances as well as in firm-by-firm time-series regressions. Regressions with hourly data produce similar results. The positivity of cross-market lambdas implies that information that arrives to the market via order flows is primarily about asset means. Moreover, the relative magnitude of the within-market price impacts is larger than the cross-market price impacts in the data, as predicted by multiasset Kyle models. We also show that bond and stock markets are integrated in their responses to order flows: the components of returns explained by order flows are highly correlated across markets, even though the order flows themselves are virtually uncorrelated, and neither leads the other.

It is possible that there is more evidence of information about risks in other markets. Rourke (2014) documents that put and call orders have significant positive effects on straddle returns, which suggests that option order flow may contain information about risks. It is easy to generalize our model to include options in addition to debt and equity; however, the binary state assumption may be too simple for analyzing additional markets. With only two states, if information is purely about the asset risk, then the stock, call, and put values should be monotonically related, all being higher in the high-risk state. This implies that cross-market put/call lambdas should be positive (they would have the sign of $\Delta P \Delta C$, where Δ denotes the state-2 value minus the state-1 value, and P and C denote the put and call, respectively). However, Rourke finds that cross-market call/put lambdas are negative. On the other hand, in a two-state world with information purely about the asset mean, the call and put values are inversely related, implying that the cross-market straddle/call and straddle/put lambdas should have opposite signs, which is again inconsistent with Rourke's results. Development and estimation of a more general model for the simultaneous analysis of debt, equity, and option markets may be a desirable objective for future research.

Appendix A. Proofs

Let $\tilde{\xi}$ denote the binary signal observed by the informed trader (for example, $\tilde{\xi}=(\tilde{\mu},\tilde{\sigma})$ in the lognormal example). The signal $\tilde{\xi}$ has two possible values, ξ_1 and ξ_2. It is convenient to write $B(\tilde{\xi})$ and $S(\tilde{\xi})$ for the expected bond and stock values conditional on the signal, so the symbols $B(\xi_i)$ and $S(\xi_i)$ replace what were denoted in the text as B_i and S_i.

Let $\mathbb{G} \stackrel{\text{def}}{=} \{\mathcal{G}_t \mid 0 \le t \le T\}$ denote the completion of the filtration generated by Z, form the enlarged filtration with σ–fields $\mathcal{G}_t \vee \sigma(\tilde{\xi})$, and let $\mathbb{F} \stackrel{\text{def}}{=} \{\mathcal{F}_t \mid 0 \le t \le T\}$ denote the completion of the enlarged filtration. The filtration $\mathbb{F}$ represents the informed trader's information.

Before beginning the proof of the proposition, we will explain why the informed trader can infer enough about the liquidity trade process Z from equilibrium prices to implement the equilibrium trading strategy. As remarked in Section 2, the equilibrium prices reveal $A_t - \alpha = \Delta' Y_t = \Delta'(X_t + Z_t)$. Define $W_t = \Delta' Z_t$. We can project Z^b and Z^s on W as

$$Z^b = \frac{\delta_b \sigma_b}{\phi} W + \sigma_b \sqrt{1-\delta_b^2}\, W^b, \tag{A1a}$$

$$Z^s = \frac{\delta_s \sigma_s}{\phi} W + \sigma_s \sqrt{1-\delta_s^2}\, W^s, \tag{A1b}$$

where

$$\delta_b = \frac{\sigma_b \Delta_B + \rho \sigma_s \Delta_S}{\phi},$$

$$\delta_s = \frac{\sigma_s \Delta_S + \rho \sigma_b \Delta_B}{\phi},$$

and where W^b and W^s are standard Brownian motions that are each independent of W (to verify, just regard Equation (A1) as a definition of W^b and W^s and check that they have unit standard deviation per unit of time and are uncorrelated with W). The martingale property of W^b and its independence from W imply that

$$\mathsf{E}[Z_T^b - Z_t^b \mid Z_t^b, Z_t^s, W_t^b, W_T \ge 0] = \frac{\delta_b \sigma_b}{\phi} \mathsf{E}[W_T - W_t \mid Z_t^b, Z_t^s, W_T \ge 0].$$

Therefore, by iterated expectations,

$$\mathsf{E}[Z_T^b - Z_t^b \mid Z_t^b, Z_t^s, W_T \ge 0] = \frac{\delta_b \sigma_b}{\phi} \mathsf{E}[W_T - W_t \mid Z_t^b, Z_t^s, W_T \ge 0].$$

Using the fact that the distribution of W_T conditional on Z_t^b and Z_t^s depends only on $W_t = \Delta' Z_t$, we have that

$$\mathsf{E}[Z_T^b - Z_t^b \mid Z_t^b, Z_t^s, W_T \ge 0]$$

is a function of W_t. Likewise,

$$\mathsf{E}[Z_T^s - Z_t^s \mid Z_t^b, Z_t^s, W_T \ge 0]$$

and these same expectations conditional on $W_T \le 0$ are functions of W_t. It follows that the equilibrium trading strategies are functions of A_t.

Now, to prove the proposition, set $R_1 = \{y \in \mathbb{R}^2 \mid \Delta' y < -\alpha\}$ and $R_2 = \{y \in \mathbb{R}^2 \mid \Delta' y > -\alpha\}$. Define $\pi^b(y) = B(\xi_1) + \Delta_B 1_{R_2}(y)$ and $\pi^s(y) = S(\xi_1) + \Delta_S 1_{R_2}(y)$. For $t \le T$, set

$$p^b(t,y) = \mathrm{e}^{-r(T-t)} \mathsf{E}\left[\pi^b(Z_T) \mid Z_t = y\right], \tag{A2a}$$

$$p^s(t,y) = \mathrm{e}^{-r(T-t)} \mathsf{E}\left[\pi^s(Z_T) \mid Z_t = y\right]. \tag{A2b}$$

We will first show that the trading strategy defined by Equation (7) is optimal for the informed trader, when prices are defined by Equation (A2). For each $y \in \mathbb{R}^2$, define $J(T,y,\xi_1)=1_{R_2}(y)\Delta' y$ and $J(T,y,\xi_2)=-1_{R_1}(y)\Delta' y$. Note that $J(T,y,\xi_i)>0$ for $y\in R_j$ and $j\neq i$. Also, for $y\in R_1\cup R_2$,

$$\frac{\partial J(T,y,\xi_i)}{\partial y^b}=\pi^b(y)-B(\xi_i), \tag{A3a}$$

$$\frac{\partial J(T,y,\xi_i)}{\partial y^s}=\pi^s(y)-S(\xi_i). \tag{A3b}$$

For $t<T$, set $J(t,y,\xi_i)=\mathsf{E}[J(T,Z_T,\xi_i)\,|\,Z_t=y]$. Note that the definition of $J(T,y,\xi_i)$ for $y\notin R_1\cup R_2$ is irrelevant for the definition of $J(t,y,\xi_i)$, because $Z_T\in R_1\cup R_2$ with probability one. We can interchange differentiation and expectation and use Equation (A3) to compute the gradient of $J(t,\cdot,\xi_i)$ as

$$\frac{\partial J(t,y,\xi_i)}{\partial y^b}=\mathsf{E}[\pi^b(Z_T)\,|\,Z_t=y]-B(\xi_i), \tag{A4a}$$

$$\frac{\partial J(t,y,\xi_i)}{\partial y^s}=\mathsf{E}[\pi^s(Z_T)\,|\,Z_t=y]-S(\xi_i). \tag{A4b}$$

Also, note that, because $J(t,Z_t,\tilde\xi)$ is by definition an $\mathbb{F}$–martingale, its drift on the filtration $\mathbb{F}$ is zero. This yields

$$J_t+\frac{1}{2}\frac{\partial^2 J}{\partial z^b}\sigma_b^2+\frac{1}{2}\frac{\partial^2 J}{\partial z^s}\sigma_s^2+\frac{\partial^2 J}{\partial z^b\partial z^s}\rho\sigma_b\sigma_s=0. \tag{A5}$$

Consider an arbitrary strategy θ for the informed trader and apply Itô's formula to $J(t,Y_t,\tilde\xi)$. Equation (A5) states that the second-order terms in $\mathrm{d}J$ cancel with $J_t\,\mathrm{d}t$, leaving

$$\mathrm{d}J(t,Y_t,\tilde\xi)=\frac{\partial J(t,Y_t,\tilde\xi)}{\partial y^b}\mathrm{d}Y_t^b+\frac{\partial J(t,Y_t,\tilde\xi)}{\partial y^s}\mathrm{d}Y_t^s.$$

Equations (A2) and (A4) imply

$$\frac{\partial J(t,Y_t,\tilde\xi)}{\partial y^b}=\mathrm{e}^{r(T-t)}p^b(t,Y_t)-B(\tilde\xi),$$

$$\frac{\partial J(t,Y_t,\tilde\xi)}{\partial y^s}=\mathrm{e}^{r(T-t)}p^s(t,Y_t)-S(\tilde\xi).$$

Therefore,

$$\begin{aligned}J(T,Y_T,\tilde\xi)&=J(0,0,\tilde x)+\int_0^T \mathrm{d}J(t,Y_t,\tilde x)\\&=J(0,0,\tilde\xi)+\int_0^T\left(\mathrm{e}^{r(T-t)}p^b(t,Y_t)-B(\tilde\xi)\right)\mathrm{d}Y_t^b\\&\quad+\int_0^T\left(\mathrm{e}^{r(T-t)}p^s(t,Y_t)-S(\tilde\xi)\right)\mathrm{d}Y_t^s.\end{aligned}$$

Recalling that $\mathrm{d}Y=\theta\,\mathrm{d}t+\mathrm{d}Z$, we see that the negative of the profit in expression (3) constitutes part of the right-hand side of this equation. Rearranging, we see that the expected profit (3) equals

$$\begin{aligned}&\mathsf{E}\left[J(0,0,\tilde\xi)-J(T,Y_T,\tilde\xi)\right]\\&\quad+\mathsf{E}\left[\int_0^T\left(\mathrm{e}^{r(T-t)}p^b(t,Y_t)-B(\tilde\xi)\right)\mathrm{d}Z_t^b+\int_0^T\left(\mathrm{e}^{r(T-t)}p^s(t,Y_t)-S(\tilde\xi)\right)\mathrm{d}Z_t^s\right].\end{aligned}$$

The stochastic integrals in this expression have zero expectations, due to the "no doubling strategies" condition. Moreover, $J(T,Y_T,\tilde\xi)\geq 0$, so the expected profit is bounded above by

$\mathsf{E}[J(0,0,\xi)]$, with the bound being achieved if and only if $J(T,Y_T,\xi)=0$ with probability 1. Thus, a trading strategy is optimal if and only if it implies $Y_T \in R_i$ whenever $\xi=\xi_i$.

When prices are given by Equation (A2), the optimality of the trading strategy defined by Equation (7) is now implied by the following, the proof of which is momentarily deferred.

Lemma 1. The following are true for the trading strategy q defined by Equation (7):

(A) There is a unique strong solution to the stochastic differential equation (SDE)

$$\mathrm{d}Y_t = q(t,Y_t,\xi)\,\mathrm{d}t + \mathrm{d}Z_t \tag{A6}$$

on the filtration $\mathbb{F}$ with initial condition $Y_0=0$.

(B) The solution Y to the SDE (A6) satisfies, with probability 1, $\xi=\xi_i \Leftrightarrow Y_T \in R_i$.

(C) The solution Y to the SDE (A6) is a Brownian motion on its own filtration with zero drift and instantaneous covariance matrix Σ (that is, it has the same law as Z).

Now, we need to show that the prices $P_t^b \stackrel{\text{def}}{=} p^b(t,Y_t)$ and $P_t^s \stackrel{\text{def}}{=} p^s(t,Y_t)$ with p^b and p^s defined by Equation (A2) satisfy Equation (2) and that Equation (6) holds, when Equation (7) defines the trading strategy. In the more formal notation of this appendix, the expression π_t appearing in Equations (2) and (6) is given by $\mathbb{P}(\xi=\xi_2 \mid \mathcal{G}_t)$. From Part B of the lemma,

$$\mathbb{P}(\xi=\xi_2 \mid \mathcal{G}_t) = \mathbb{P}(Y_T \in R_2 \mid \mathcal{G}_t).$$

Part C of the lemma implies Y is Markov on $(\Omega,\mathbb{G},\mathbb{P})$, so

$$\mathbb{P}(Y_T \in R_2 \mid \mathcal{G}_t) = \mathbb{P}(Y_T \in R_2 \mid Y_t).$$

The definition (A2) and Part C of the lemma imply

$$\begin{aligned} p^b(t,y) &= B(\xi_1) + \Delta_B \mathbb{P}(Z_T \in R_2 \mid Z_t = y) \\ &= B(\xi_1) + \Delta_B \mathbb{P}(Y_T \in R_2 \mid Y_t = y), \end{aligned}$$

Therefore,

$$P_t^b = B(\xi_1) + \Delta_B \mathbb{P}(Y_T \in R_2 \mid Y_t).$$

By the same reasoning,

$$P_t^s = S(\xi_1) + \Delta_S \mathbb{P}(Y_T \in R_2 \mid Y_t).$$

This verifies Equation (2). Finally, the definition of R_2 and Part C of the lemma imply

$$\mathbb{P}(Y_T \in R_2 \mid Y_t) = \mathrm{N}\left(\frac{\alpha + \Delta' Y_t}{\phi\sqrt{T-t}}\right),$$

which verifies Equation (6).

To establish the remaining claims in the proposition, first apply Itô's formula to Equation (2) to obtain

$$\mathrm{d}P_t = r P_t\,\mathrm{d}t + \mathrm{e}^{-r(T-t)} \Delta\,\mathrm{d}\pi_t.$$

Using Equation (6) for π_t and Itô's formula again yields

$$\mathrm{d}\pi_t = \frac{1}{\phi\sqrt{T-t}}\,\mathrm{n}\left(\frac{\alpha+\Delta' Y_t}{\phi\sqrt{(T-t)}}\right)\Delta'\mathrm{d}Y_t.$$

Combining these produces $\mathrm{d}P_t = r P_t\,\mathrm{d}t + \Lambda_t\,\mathrm{d}Y_t$ with $\Lambda_t = \kappa(t,\alpha+\Delta' Y_t)\Delta\Delta'$ as claimed. Part C of the lemma shows that the aggregate order process Y is a Brownian motion with zero drift and instantaneous covariance matrix Σ given the market makers' information. Part B of the lemma states that $\cup_i \{\omega \mid \xi(\omega)=\xi_i, Y_T(\omega)\in R_i\}$ is a set of states of the world that has probability 1. Consider a state ω such that $\xi(\omega)=\xi_1$ and $Y_T(\omega)\in R_1$. Then $-[\alpha+\Delta' Y_T(\omega)]>0$. Set $\epsilon(\omega)=-[\alpha+\Delta' Y_T(\omega)]/2>0$. By the path continuity of Y, there exists $\tau(\omega)<T$ such that $\Delta' Y_t(\omega)-\Delta' Y_T(\omega)<\epsilon(\omega)$ for all $t>\tau(\omega)$, so

$$\begin{aligned}\alpha+\Delta' Y_t(\omega)+\epsilon(\omega) &= \alpha+\Delta' Y_T(\omega)+\epsilon(\omega)+[\Delta' Y_t(\omega)-\Delta' Y_T(\omega)]\\ &< \alpha+\Delta' Y_T(\omega)+2\epsilon(\omega)\\ &= 0\end{aligned}$$

for $t>\tau(\omega)$. Therefore, $\alpha+\Delta' Y_t(\omega)<-\epsilon(\omega)$ for $t>\tau(\omega)$. For such t, Equation (6) for π_t implies

$$\pi_t(\omega)=\mathrm{N}\left(\frac{\alpha+\Delta' Y_t(\omega)}{\phi\sqrt{T-t}}\right)<\mathrm{N}\left(\frac{-\epsilon(\omega)}{\phi\sqrt{T-t}}\right).$$

The right-hand side of this inequality converges to zero as $t\to T$, so $\pi_t(\omega)\to 0$. A similar argument shows that $\pi_t(\omega)\to 1$ for each ω such that $\tilde{x}(\omega)=\xi_2$ and $Y_T(\omega)\in R_2$. It now follows immediately from Equation (2) that there is convergence to strong-form efficiency as claimed in the proposition.

Proof. To put the stochastic differential equation (A6) in a more standard form, define the process $\hat{Y}_t=(Y_t,\xi)$ with random initial condition $\hat{Y}_0=(0,\xi)$, and augment the SDE (A6) with the equation $\mathrm{d}\xi=0$. The existence of a unique strong solution $\hat{Y}$ to this enlarged system follows from Lipschitz and growth conditions satisfied by q. See Karatzas and Shreve (1988, Theorem 5.2.9).

The uniqueness in distribution of weak solutions of stochastic differential equations (Karatzas and Shreve, 1988, Theorem 5.3.10) implies that we can demonstrate Properties B and C by exhibiting a weak solution for which they hold. Let $\mathbb{P}$ denote the probability measure. Let 1_A denote the zero-one indicator function of a set A. Directly from the definitions of π_0, α, R_i and ξ, we have

$$\mathbb{P}(\xi=\xi_2)=\pi_0=\mathrm{N}\left(\frac{\alpha}{\phi\sqrt{T}}\right)=\mathbb{P}(Z_T\in R_2). \tag{A7}$$

Define

$$M_T=\frac{1}{1-\pi_0}\left[1_{R_1}(Z_T)1_{\{\xi_1\}}(\xi)\right]+\frac{1}{\pi_0}\left[1_{R_2}(Z_T)1_{\{\xi_2\}}(\xi)\right].$$

The $\mathbb{P}$–independence of Z_T and ξ and (A7) imply

$$\mathsf{E}[M_T]=\frac{\mathbb{P}(Z_T\in R_1)\mathbb{P}(\xi=\xi_1)}{1-\pi_0}+\frac{\mathbb{P}(Z_T\in R_2)\mathbb{P}(\xi=\xi_2)}{\pi_0}=1.$$

Define a new measure $\mathbb{Q}$ on $(\Omega,\mathcal{F}_T)$ by $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}=M_T$. Define $H_i(t,y)=\mathbb{P}(Z_T\in R_i\mid Z_t=y)$ for $i=1,2$, and for $t\le T$, set

$$M_t=\mathsf{E}[M_T\mid\mathcal{F}_t]=\frac{1}{1-\pi_0}\left[H_1(t,Z_t)1_{\{\xi_1\}}(\xi)\right]+\frac{1}{\pi_0}\left[H_2(t,Z_t)1_{\{\xi_2\}}(\xi)\right]. \tag{A8}$$

Obviously, only one of the terms on the right-hand side of the definition (A8) is nonzero in any state of the world. Thus,

$$\frac{\mathrm{d}M_t}{M_t}=\frac{\mathrm{d}H_1(t,Z_t)}{H_1(t,Z_t)}1_{\{\xi_1\}}(\xi)+\frac{\mathrm{d}H_2(t,Z_t)}{H_2(t,Z_t)}1_{\{\xi_2\}}(\xi). \tag{A9}$$

From Itô's formula, and the fact that $H_i(t,Z_T)$ is a $\mathbb{P}$–martingale, we obtain

$$\mathrm{d}H_i(t,Z_t)=\nabla H_i(t,Z_t)'\mathrm{d}Z_t, \tag{A10}$$

where ∇ denotes the gradient with respect to the Z_t argument. Let $f(\cdot \mid t, y)$ denote the density function of Z_T conditional on $Z_t = y$. Let ∇_y denote a gradient with respect to $y \in \mathbb{R}^2$. Then,

$$\begin{aligned}\nabla_y H_i(t,y) &= \nabla_y \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{R_i}(z) f(z \mid t,y)\,\mathrm{d}z^b\,\mathrm{d}z^s \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{R_i}(z)\nabla_y f(z \mid t,y)\,\mathrm{d}z^b\,\mathrm{d}z^s \\ &= \frac{1}{T-t}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{R_i}(z)\Sigma^{-1}(z-y) f(z \mid t,y)\,\mathrm{d}z^b\,\mathrm{d}z^s.\end{aligned}$$

It follows that

$$\begin{aligned}\frac{1}{H_i(t,y)}\nabla_y H_i(t,y) &= \frac{1}{T-t}\Sigma^{-1}\mathsf{E}[Z_T - y \mid Z_t = y, Z_T \in R_i] \\ &= \Sigma^{-1} q(t,y,\xi_i). \qquad \text{(A11)}\end{aligned}$$

Thus, Equation (A10) implies

$$\frac{\mathrm{d}H_t(t,Z_t)}{H_i(t,Z_t)} = q(t,Z_t,\xi_i)'\Sigma^{-1}\mathrm{d}Z_t,$$

and from Equation (A9) we obtain

$$\frac{\mathrm{d}M_t}{M_t} = q(t,Z_t,\tilde{\xi})'\Sigma^{-1}\mathrm{d}Z_t.$$

Therefore,

$$\left(\frac{\mathrm{d}M_t}{M_t}\right)\mathrm{d}Z = q(t,Z_t,\tilde{\xi})\mathrm{d}t.$$

Consequently, by Girsanov's Theorem, Z^* defined by $Z_0^* = 0$ and

$$\mathrm{d}Z_t^* = -q(t,Z_t,\tilde{\xi})\mathrm{d}t + \mathrm{d}Z_t \qquad \text{(A12)}$$

is a Brownian motion (with zero drift and instantaneous covariance matrix Σ) on the filtration $\mathbb{F}$ relative to $\mathbb{Q}$. Equation (A12) implies that Z is a weak solution of the SDE (A6) relative to the Brownian motion Z^* on the filtered probability space $(\Omega, \mathbb{F}, \mathbb{Q})$. We want to establish Properties B and C for this weak solution.

Property B is equivalent to the condition that

$$\sum_{i=1}^{2}\mathsf{E}\left[1_{R_i}(Y_T)1_{\{\xi_i\}}(\tilde{\xi})\right] = 1.$$

It holds for the weak solution Z on $(\Omega, \mathbb{F}, \mathbb{Q})$ iff

$$\sum_{i=1}^{2}\mathsf{E}^{\mathbb{Q}}\left[1_{R_i}(Z_T)1_{\{\xi_i\}}(\tilde{\xi})\right] = 1,$$

where $\mathsf{E}^{\mathbb{Q}}$ denotes expectation with respect to $\mathbb{Q}$. This is equivalent to

$$\sum_{i=1}^{2}\mathsf{E}\left[M_T 1_{R_i}(Z_T)1_{\{\xi_i\}}(\tilde{\xi})\right] = 1.$$

By the definition of M_T,

$$\sum_{i=1}^{2} M_T 1_{R_i}(Z_T)1_{\{\xi_i\}}(\tilde{\xi}) = M_T,$$

and $\mathsf{E}[M_T] = 1$, so Property B holds.

It remains to establish Property C for this weak solution, meaning that Z is a Brownian motion on $(\Omega, \mathbb{G}, \mathbb{Q})$. Because Z is a Brownian motion on $(\Omega, \mathbb{G}, \mathbb{P})$, it suffices to show that $\mathbb{Q} = \mathbb{P}$ when both are restricted to $\mathcal{G}_T$. Because $\mathbb{G}$ is left-continuous, it actually suffices to show that $\mathbb{Q} = \mathbb{P}$ when both are restricted to $\mathcal{G}_\tau$ for arbitrary $\tau < T$. This holds if for all $t_1 < \cdots < t_n \leq \tau$ and all Borel B we have

$$\mathbb{P}((Z_{t_1}, \ldots, Z_{t_n}) \in B) = \mathbb{Q}((Z_{t_1}, \ldots, Z_{t_n}) \in B).$$

The right-hand side of this equals

$$\begin{aligned} \mathsf{E}[M_T 1_B(Z_{t_1}, \ldots, Z_{t_n})] &= \frac{1}{1-\pi_0} \mathsf{E}\left[1_{R_1}(Z_T) 1_{\{\xi_1\}}(\tilde{\xi}) 1_B(Z_{t_1}, \ldots, Z_{t_n})\right] \\ &\quad + \frac{1}{\pi_0} \mathsf{E}\left[1_{R_2}(Z_T) 1_{\{\xi_2\}}(\tilde{\xi}) 1_B(Z_{t_1}, \ldots, Z_{t_n})\right] \\ &= \mathsf{E}\left[1_{R_1}(Z_T) 1_B(Z_{t_1}, \ldots, Z_{t_n})\right] + \mathsf{E}\left[1_{R_2}(Z_T) 1_B(Z_{t_1}, \ldots, Z_{t_n})\right] \\ &= \mathsf{E}\left[1_{R_1 \cup R_2}(Z_T) 1_B(Z_{t_1}, \ldots, Z_{t_n})\right] \\ &= \mathbb{P}((Z_{t_1}, \ldots, Z_{t_n}) \in B), \end{aligned}$$

using the $\mathbb{P}$–independence of Z and $\tilde{\xi}$ for the second equality. ■

Appendix B. Data

The empirical tests draw data from several databases. Here we detail the sample selection process, data filters, and methods for linking the databases.

TRACE: TRACE contains transactions data for corporate bonds. We employ several filters to ensure accuracy of the TRACE data. Following Dick-Nielsen (2009), we first delete any duplicate observations by identifying same-bond entries with the same intraday message sequence number (every report in TRACE has a unique message sequence number within a day). Second, we delete any observations that have been reversed. This involves deleting the reversal report and the original report that the reversal report references. Third, we delete any original report that has been corrected later in the trading day.

When calculating bonds returns, we employ several price filters to minimize the effect of price errors. First, we require that prices satisfy:

$$|P - \text{med}(P, 20)| \leq 5 \times \text{MAD}(P, 20) + \$1$$

where P is the transaction price, $\text{med}(P, 20)$ is the centered rolling median over 20 price observations, and $\text{MAD}(P, 20)$ is the median absolute deviation of the price.[16] We also require that the reported price be between \$1 and \$500, and we delete an observation if its price is not within 20% of the bond's median price for that day or is not within 20% of the price from the previous transaction.[17] These filters remove approximately 2.5% of bond trade observations.

FISD: We match bond transactions data from TRACE to bond attributes in FISD using the bond's nine-digit CUSIP number.

CRSP: Bonds are initially matched to issuer equity information using the six-digit issuer CUSIP and the CRSP stocknames file. For any unmatched bonds, we use the issuer CUSIP associated with the parent ID in FISD to match to CRSP. On some days on which equities trade, bond markets are

16 This is the filter used by Rossi (2009).

17 This is the filter used by Han and Zhou (2008).

closed or have significantly reduced volumes and numbers of bonds traded, particularly around holidays. We eliminate these days (about five per year).

TAQ: Issuers identified in CRSP are matched to intraday stock transactions in TAQ each month using the eight-digit equity CUSIP and the monthly TAQ master files. We use trades and quotes occurring during the trading day and exclude observations with price or volume values of zero. In applying the Lee-Ready algorithm, trades are matched to quotes in the same second.

References

Acharya, V. V., and T. C. Johnson. 2007. Insider trading in credit derivatives. *Journal of Financial Economics* 84:110–41.

———. 2010. More insiders, more insider trading: Evidence from private-equity buyouts. *Journal of Financial Economics* 98:500–23.

Alexander, G. J., A. K. Edwards, and M. G. Ferri. 2000. What does Nasdaq's high-yield bond market reveal about bondholder-stockholder conflicts? *Financial Management* 29:23–39.

Andrade, S. C., C. Chang, and M. S. Seasholes. 2008. Trading imbalances, predictable reversals, and cross-stock price pressure. *Journal of Financial Economics* 88:406–23.

Back, K. 1992. Insider trading in continuous time. *Review of Financial Studies* 5:387–409.

———. 1993. Asymmetric information and options. *Review of Financial Studies* 6:435–72.

Back, K., K. Crotty, and T. Li. 2014. Can information asymmetry be identified from order flows alone? Working Paper, Rice University and City University of Hong Kong.

Bai, J., and P. Collin-Dufresne. 2013. The determinants of the CDS-bond basis during the financial crisis of 2007–2009. Working Paper, Georgetown University and Swiss Finance Institute.

Bessembinder, H., W. Maxwell, and K. Venkataraman. 2006. Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. *Journal of Financial Economics* 82:251–88.

Bodnaruk, A., and M. Rossi. 2013. Dual ownership, returns, and voting in mergers. Working Paper, University of Notre Dame and Texas A&M University.

Boot, A., and A. Thakor. 1993. Security design. *Journal of Finance* 48:1349–78.

Boulatov, A., T. Hendershott, and D. Livdan. 2013. Informed trading and portfolio returns. *Review of Economic Studies* 80:35–72.

Breen, W. J., L. S. Hodrick, and R. A. Korajczyk. 2002. Predicting equity liquidity. *Management Science* 48: 470–83.

Caballé, J., and M. Krishnan. 1994. Imperfect competition in a multi-security market with risk neutrality. *Econometrica* 62:695–704.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2011. Robust inference with multiway clustering. *Journal of Business and Economic Statistics* 29:238–49.

Campbell, J. Y., and G. B. Taksler. 2003. Equity volatility and corporate bond yields. *Journal of Finance* 58: 2321–49.

Chan, K., Y. P. Chung, and W.-M. Fong. 2002. The informational role of stock and option volume. *Review of Financial Studies* 15:1049–75.

Chan, K., A. Menkveld, and Z. Yang. 2007. The informativeness of domestic and foreign investors' stock trades: Evidence from the perfectly segmented chinese market. *Journal of Financial Markets* 10:391–415.

Chang, C., and X. Yu. 2010. Informational efficiency and liquidity premium as the determinants of capital structure. *Journal of Financial and Quantitative Analysis* 45:401–40.

Chordia, T., A. Sarkar, and A. Subrahmanyam. 2005. An empirical analysis of stock and bond market liquidity. *Review of Financial Studies* 18:85–129.

Collin-Dufresne, P., R. S. Goldstein, and J. S. Martin. 2001. The determinants of credit spread changes. *Journal of Finance* 56:2177–207.

Cremers, M., J. Driessen, P. Maenhout, and D. Weinbaum. 2008. Individual stock-option prices and credit spreads. *Journal of Banking and Finance* 32:2706–15.

Dick-Nielsen, J. 2009. Liquidity biases in TRACE. *Journal of Fixed Income* 19:43–55.

Downing, C., S. Underwood, and Y. Xing. 2009. The relative informational efficiency of stocks and bonds: An intraday analysis. *Journal of Financial and Quantitative Analysis* 44:1081–102.

Duffie, D., N. Garleanu, and L. H. Pedersen. 2005. Over-the-counter markets. *Econometrica* 73:1815–47.

Edwards, A. K., L. E. Harris, and M. S. Piwowar. 2007. Corporate bond market transaction costs and transparency. *Journal of Finance* 62:1421–51.

Feldhütter, P. 2012. The same bond at different prices: Identifying search frictions and selling pressures. *Review of Financial Studies* 25:1155–206.

Fulghieri, P., and D. Lukin. 2001. Information production, dilution costs, and optimal security design. *Journal of Financial Economics* 61:3–42.

Goldstein, M. A., E. S. Hotchkiss, and E. R. Sirri. 2007. Transparency and liquidity: A controlled experiment on corporate bonds. *Review of Financial Studies* 20:235–73.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka. 2009. Do liquidity measures measure liquidity? *Journal of Financial Economics* 92:153–81.

Green, R. C., B. Hollifield, and N. Schürhoff. 2007. Financial intermediation and the costs of trading in an opaque market. *Review of Financial Studies* 20:275–314.

Han, S., and H. Zhou. 2008. Effects of liquidity on the nondefault component of corporate yield spreads: Evidence from intraday transactions data. Working Paper, Federal Reserve Board.

Han, S., and X. Zhou. 2014. Informed bond trading, corporate yield spreads, and corporate default prediction. *Management Science* 60:675–94.

Hasbrouck, J. 1991. Measuring the information content of stock trades. *Journal of Finance* 46:179–207.

———. 1995. One security, many markets: Determining the contributions to price discovery. *Journal of Finance* 50:1175–99.

———. 2009. Trading costs and returns for U.S. equities: Estimating effective costs from daily data. *Journal of Finance* 64:1445–77.

He, Z., and K. Milbradt. 2014. Endogenous liquidity and defaultable bonds. *Econometrica* 82:1443–508.

Hilscher, J., J. M. Pollet, and M. Wilson. forthcoming. Are credit default swaps a sideshow? Evidence that information flows from equity to CDS markets. *Journal of Financial and Quantitative Analysis.*

Hotchkiss, E. S., and T. Ronen. 2002. The informational efficiency of the corporate bond market: An intraday analysis. *Review of Financial Studies* 15:1325–54.

Jones, C. M., G. Kaul, and M. L. Lipson. 1994. Transactions, volume, and volatility. *Review of Financial Studies* 7:631–51.

Kapadia, N., and X. Pu. 2012. Limited arbitrage between equity and credit markets. *Journal of Financial Economics* 105:542–64.

Karatzas, I., and S. E. Shreve. 1988. *Brownian motion and stochastic calculus*. New York: Springer-Verlag.

Kedia, S., and X. Zhou. 2014. Informed trading around acquisitions: Evidence from corporate bonds. *Journal of Financial Markets* 18:182–205.

Korajczyk, R. A., D. J. Lucas, and R. L. McDonald. 1991. The effect of information releases on the pricing and timing of equity issues. *Review of Financial Studies* 4:685–708.

Kwan, S. H. 1996. Firm-specific information and the correlation between individual stocks and bonds. *Journal of Financial Economics* 40:63–80.

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315–36.

Lee, C. M., and M. J. Ready. 1991. Inferring trade direction from intraday data. *Journal of Finance* 46:733–46.

Lesmond, D. A., P. F. O'Connor, and L. W. Senbet. 2008. Capital structure and equity liquidity. Working Paper, University of Maryland.

Mao, Y. 2013. Price discovery in the stock and corporate bond markets. Working Paper, Indiana University.

McCracken, J. 2008. Delphi recovery hits snag; bankruptcy-court exit sidetracked by disputes with investor groups. *Wall Street Journal,* March 13.

O'Hara, M., C. Yao, and M. Ye. 2014. What's not there: Odd lots and market data. *Journal of Finance* 69: 2199–236.

Pasquariello, P., and C. Vega. forthcoming. Strategic cross-trading in the U.S. stock market. *Review of Finance*.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22:435–80.

Ronen, T., and X. Zhou. 2013. Trade and information in the corporate bond market. *Journal of Financial Markets* 16:61–103.

Rossi, M. 2009. Realized volatility, liquidity, and corporate yield spreads. Working Paper, University of Notre Dame.

Rourke, T. 2014. The delta- and vega-related information content of near-the-money option market trading activity. *Journal of Financial Markets* 20:175–93.

Schultz, P. 2001. Corporate bond trading costs: A peek behind the curtain. *Journal of Finance* 56:677–98.

Szockyj, E. 1993. Insider trading: The SEC meets Carl Karcher. *Annals of the American Academy of Political and Social Science* 52:46–58.

Thompson, S. B. 2011. Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics* 99:1–10.

Tookes, H. E. 2008. Information, trading, and product market interactions: Cross-sectional implications of informed trading. *Journal of Finance* 63:379–413.

Veronesi, P. 1999. Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies* 12:975–1007.

Wei, J., and X. Zhou. 2012. Informed trading in corporate bonds prior to earnings announcements. Working Paper, University of Toronto and Rutgers University.

Zhu, H. 2012. Finding a good price in opaque over-the-counter markets. *Review of Financial Studies* 25: 1255–85.