The Complexity Yield Puzzle: A Textual Analysis of Municipal Bond Disclosures

Michael Farrell, Dermot Murphy, Marcus Painter, Guangli Zhang*

July 11, 2023

Abstract

We study the effect of official statement textual complexity on municipal borrowing costs (yield). Theoretically, complexity increases yield if the marginal cost of complexity to unsophisticated traders exceeds the marginal value to sophisticated traders. Empirically, we provide evidence of an economically significant complexity-yield premium that is especially large for bonds that have greater tax-induced exogenous demand from unsophisticated investors, less complementary information from ratings agencies, or greater risk. Complexity also increases yield volatility and the markup differential between retail and institutional investors. We decompose complexity into multiple topics and find that "legal complexity" matters most for the complexity-yield premium. Despite this premium, complexity has been trending upward across all credit categories. We show that this trend is attributable to increased regulatory enforcement from the SEC.

JEL classification: G12, G14, G18, H74

Keywords: municipal bonds, official statements, complexity, textual analysis

*Michael Farrell: Department of Finance, University of Wisconsin–Milwaukee, Milwaukee, WI, 53211; e-mail: farrell1@uwm.edu. Dermot Murphy: University of Illinois at Chicago, College of Business, Department of Finance, 601 South Morgan Street, Chicago, IL, 60607; e-mail: murphyd@uic.edu. Marcus Painter: Department of Finance, Saint Louis University, Saint Louis, MO, 63103; e-mail: marc.painter@slu.edu. Guangli Zhang: Department of Finance, Saint Louis University, Saint Louis, MO, 63103; e-mail: guangli.zhang@slu.edu. We thank Dan Garrett, Igor Cunha, and seminar participants at CUNY-Queens, Oklahoma State University, Saint Louis University, Southern Illinois University, the Virtual Municipal Finance Workshop, and the 2022 Dayton Summer Finance Workshop for helpful comments. All errors are our own.

1 Introduction

The transmission of information is a joint activity of the sender and the receiver. In financial markets, the sender discloses information that is relevant to the security, and the receiver commits time and resources to process that information. In order to effectively communicate the relevant risks, the sender needs to account for the information processing ability of their investor clientele. Institutional investors may extract more value from complex disclosures because they have the resources to process the associated information, while retail investors may extract more value from simplified disclosures due to a relative lack of resources. As a result, there can be information asymmetries between investor types, even for public disclosures (Fishman and Hagerty, 2003; Veldkamp, 2011). In this study, we examine how disclosure complexity affects the cost of capital and market quality, and how these effects vary in the cross-section of investor clientele. Given the stark increase in retail participation across asset classes in the past decade (De Sousa, 2021; WSJ, 2021; Bloomberg, 2022) and additional evidence that modern retail investors are more likely to ignore disclosures (Casey and Macey, 2020; WSJ, 2022a; Moss et al., 2023), it is more important than ever to understand how disclosure complexity affects market outcomes through the clientele channel.

Investor clientele heterogeneity is particularly stark in the $4 trillion U.S. municipal bond market, where direct participation by individual investors is high relative to other markets (Bergstresser, 2023), nearly half of outstanding bonds are held by individual investors (Bagley et al., 2023c), and there is a large degree of tax-induced ownership segmentation across states (Babina et al., 2021; Garrett et al., 2023). Therefore, the municipal bond market provides a useful laboratory for studying the clientele-driven benefits and costs of disclosure complexity because of its highly varied investor clientele in addition to its lack of common structure for the accompanying official statements. We take the first comprehensive look at the textual complexity of over 80,000 municipal bond official statements to shed light on how municipal disclosure complexity affects local governments' cost of capital and secondary market quality in the cross-section of clientele.

We first provide strong evidence of a positive complexity-yield premium in the municipal bond market. In particular, we find that a one standard deviation increase in official statement complexity (a composite measure which is broadly based on the length and readability of the document) is associated with a 4.9 basis point increase in the offering yield spread. In terms of economic magnitude, this effect represents 7.9% of one standard deviation in the offering yield spread, or 5.1% of the credit spread between AAA and BBB municipal bonds. For an issue with an average size of $32 million and an average duration of 8 years, the effect represents a 39.2 basis point reduction in the municipal bond price, or $125,000 in additional interest payments. Complex official statements are also associated with higher price volatility and larger average markups on the secondary market. Interestingly, we find that complex official statements are associated with lower markups for institutional investors and higher markups for retail investors, suggesting that the well-documented positive markup differential between retail and institutional investors is exacerbated by official statement complexity. Overall, our baseline results indicate that official statement complexity is associated with higher borrowing costs and reduced average market quality.

The empirical findings are consistent with the central predictions of our modified Grossman and Stiglitz (1980) noisy rational expectations equilibrium (REE) model with an exogenous disclosure technology that increases signal precision for sophisticated investors and decreases signal precision for unsophisticated investors.[1] In particular, our model predicts that disclosure complexity increases the cost of capital and reduces market quality as long as the total marginal cost of disclosure complexity to unsophisticated investors exceeds the total marginal benefit to sophisticated investors. Given our empirical findings, this latter condition appears to hold. As a reality check, we plug our estimate of the complexity-yield premium (and other reasonable parameter estimates) into our theoretical model to make inferences about the overall clientele breakdown in the municipal bond market. We find that sophisticated investors represent 27% of fundamental (non-liquidity) demand for municipal bonds, with unsophisticated investors making up the remaining 73%. If we directly estimate the clientele breakdown using clientele statistics from the Securities Industry and Financial Markets Association (SIFMA), we find that 25.4% of fundamental demand comes from sophisticated investors, which is reasonably close to our model-implied estimate.[2] Therefore,

[1]See Goldstein and Yang (2017) for a comprehensive treatment of noisy REE models and their extensions which derive from seminal works by Grossman and Stiglitz (1980), Hellwig (1980), and Verrecchia (1982).

[2]In particular, using clientele statistics from SIFMA, we estimate that the breakdown of demand in the municipal bond market is 47% unsophisticated fundamental investors, 16% sophisticated fundamental investors, and 37% liquidity investors. Therefore, $16/(16+47) = 25.4\%$ of fundamental demand comes from

our empirical findings and model-implied estimates provide a reasonable approximation of the observed clientele breakdown in the municipal bond market.

A key implication of our model is that the cost (or benefit) of disclosure complexity depends on the share of investors that are sophisticated versus unsophisticated. We are able to test this prediction using multiple proxies for the presence of retail investors from the municipal finance literature.[3] Our first proxy is the tax privilege provided to investors for holding municipal bonds issued in their state of residence, which is known to increase the proportion of retail investors that are active in the local municipal bond market (Babina et al., 2021). Our second and third proxies are related to the bank qualification status and the size of the offering price premium on the bond, both of which are known to attract a larger proportion of institutional investors (Bagley et al., 2023a). Across all proxies, we find that the complexity-yield premium is driven by issuances that are more likely to have a larger retail investor base. We also find the complexity-yield premium is significantly larger for bonds issued via negotiation with an underwriter, which typically allow retail investors to participate in a priority purchase window. Overall, these findings provide identification of our proposed clientele channel for the positive complexity-yield premium.

The municipal bond market offers unique opportunities to identify instances where complexity in the official statement is more relevant for potential investors and the cost of capital. For example, official statement complexity is less likely to matter if third-party credit rating reports already provide clear external signals about the quality of the issuance. Similarly, official statement complexity is less likely to matter for general obligation bond issues because potential investors already have access to historical economic information from county government websites and other sources, unlike revenue bond issues which rely on a specific revenue source and typically do not have any historical information from external sources. The evidence we find is consistent with these conjectures. In particular, we find that the complexity-yield premium for unrated bonds (9.1 basis points) is almost twice as

sophisticated investors.

[3]Retail investors are not necessarily unsophisticated and thus incapable of processing complex information. In fact, there is evidence that retail investor trades are informed in the equity market (Kelley and Tetlock, 2013; Boehmer et al., 2021; Farrell et al., 2022). Rather, we argue that retail investors will be more constrained in terms of the time and technical resources necessary to extract precise signals of value from complex statements.

large than the average premium (4.9 basis points). Furthermore, we find no evidence of a positive complexity-yield premium for general obligation bonds, and strong evidence of sizable complexity-yield premium for revenue bonds, especially those that are unrated (14.0 basis points). Overall, this evidence indicates that the clientele-driven complexity-yield premium is exacerbated when there is a lack of complementary information from credit rating agencies or other public sources.

Given that official statement complexity is costly for issuers, one might expect to see official statements being simplified over time. However, our time-series analysis shows that official statement complexity has actually steadily increased over our sample period. In particular, we find that the average number of pages and average number of words in an official statement have increased by 40% and 64%, respectively, while the average grade level required to understand the statement has increased by about 1.8. Broadly, these changes represent a 0.50 standard deviation increase in our composite complexity measure. This upward trend in complexity also does not appear to be driven by instrumental complexity of the bonds (measured using bond characteristics such as the callability of the bond), as we see a similar trend after controlling for observable bond characteristics. We call this trend the complexity-yield puzzle, and we devote the remainder of our analysis to exploring reasons for this puzzle.

We explore two non-mutually exclusive hypotheses that may help explain the complexity-yield puzzle. The catering hypothesis posits that issuers write complex statements in an effort to attract the growing proportional base of institutional investors who have the resources to extract precise signals from complex official statements (Bergstresser and Cohen, 2016). The regulatory burden hypothesis posits that issuers increasingly add information to official statements in response to heightened regulatory oversight (Dyer et al., 2017). Our evidence primarily supports the regulatory burden hypothesis, as we find that official statements have become significantly more complex in states that experienced a relatively higher surge in local SEC enforcement officers during a major enforcement push by the SEC in the mid-2010s. There is less support for the catering hypothesis, as official statement complexity has uniformly increased across bond types, including those that are more likely to attract institutional investors such as bank-qualified bonds and high-priced bonds. Furthermore, we find that the point estimate of the complexity-yield premium is stable over our sample, suggesting that issuers have not effectively targeted institutions in order to bring down the complexity-yield premium. However, we do find that official statement complexity has increased relatively more for bonds issued in states with no local tax privilege, where institutions are more likely to comprise a larger percentage of the investor base, lending some support to the catering hypothesis. Overall, our evidence primarily supports the regulatory burden hypothesis for explaining the upward trend in complexity.

To further investigate the factors influencing the increase in complexity, we employ a machine learning technique that enables us to create structured topics for the official statements. The technique, Latent Dirichlet Allocation (LDA), has been used in prior finance settings but is especially useful for municipal bond official statements because these documents are not required to follow a consistent structured format. The LDA process allows us to narrow down what proportion of topics are discussed in each statement. We identify four primary topics – risk, legal, cash flow, and description – and measure complexity within each topic. We find that the complexity premium is primarily driven by legal complexity, which further supports our second hypothesis that complexity and the associated complexity-yield premium are tightly associated with regulatory burden.

Our study broadly contributes to the literature on information asymmetries, disclosure processing costs, and the cost of capital. In a recent literature review, Blankespoor et al. (2020) note that there are three steps to process a disclosure: awareness, acquisition, and integration. The mechanism underlying our complexity-yield premium is most related to the integration step, in that some investors have more difficulty integrating complex disclosures into their information sets, which in theory can generate information asymmetries that ultimately affect the cost of capital (Easley and O'Hara, 2004). Previous studies have shown that individual investors are less likely to invest in firms that issue complex disclosures (Miller, 2010; Lawrence, 2013) and more likely to underreact to disclosures (Ben-Raphael et al., 2017), especially when those disclosures are complex (You and Zhang, 2009) or the associated information processing costs are high (Hirshleifer et al., 2009). Our study bridges the gap between the literatures on disclosure processing costs, information asymmetries, and the cost of capital by providing model-driven evidence of a large complexity-yield premium that strongly depends on the investor clientele. This is particularly important in the public finance space and regulatory circles because a higher cost of capital is ultimately borne by the local taxpayer.

In addition, our paper is among the first to use natural language processing in the public finance setting where the economics of disclosure are unique due to tax-induced variation in local investor clientele and lack of a common structure for official statements. Textual analysis has been employed in other settings to show that investors react to the tone of firm 10-Ks (Loughran and McDonald, 2011) but under-react to changes in the text of 10-Ks (Cohen et al., 2020). Natural language processing has also been used to create measures of firm financial constraints (Bodnaruk et al., 2015), gauge the effect of the SEC's Plain English Rule on firms' disclosure style (Loughran and McDonald, 2014b), and document how 10-Ks have developed through time (Dyer et al., 2017).[4] Focusing on the readability of corporate disclosures, Ertugrul et al. (2017) show that length and ambiguity in annual reports are associated with higher loan spreads through the managerial information hoarding channel, while Bae et al. (2022) find that investors appreciate both information complexity and the richness of information content in 10-K reports. Our paper extends this literature by documenting an association between cost of capital and complex disclosure in the municipal bond market that hinges on the composition of investor clientele.

Lastly, this paper contributes to the collective knowledge on relevant determinants of pricing and trading activity in the municipal bond market. An implicit assumption in our model is that municipal investors pay attention to primary market disclosure. However, empirical evidence is mixed on whether information updates by municipalities affect trading activity in the secondary market. Earlier studies found no evidence that government annual reports influence bond prices (Ingram et al., 1989; Reck and Wilson, 2006). However, more recent evidence finds that secondary trading activity increases following timely annual disclosures by municipalities (Cuny et al., 2022), and that access to municipal disclosures reduces secondary market trading costs for individual investors (Cuny, 2018). Municipal investors also rely greatly on credit ratings to assess bond quality (Cornaggia et al., 2018), which raises the question of whether primary market disclosures are even referenced by investors. Our findings shed light on this question by showing that official statement complexity does influence

[4]See Loughran and McDonald (2020) for a recent review of other work in this area. Other related work has used machine learning and topic modelling techniques to identify risk factors that firms disclose (Lopez-Lira, 2023), the likelihood of firm misreporting (Brown et al. 2020; Hoberg and Lewis 2017), and narratives in business news (Bybee et al., 2021).

the pricing of municipal bonds, which ultimately affects local economic conditions (Adelino et al., 2017).[5] Our work also complements recent findings in Brancaccio and Kang (2022) that the number of non-standard provisions attached to a municipal bond is associated with increased trading frictions. While Brancaccio and Kang (2022) measure the complexity of bond design, our goal is to identify the complexity of bond communication while controlling for the features that make the bond design complex.[6] The growing debate on the need for transparency in municipal disclosure makes our findings relevant from a policy perspective, especially given ongoing reports that disclosure in the municipal bond market "in some ways remains stuck in the last century" (WSJ, 2022b).

The rest of this paper is organized as follows. Section 2 provides institutional background for the municipal bond market. Section 3 presents a theoretical framework that examines the impact of disclosure complexity on market outcomes. Section 4 describes the data sets and sampling restrictions, and provides summary statistics for the data used in this study. Section 5 presents our baseline results showing how complexity affects yields and other market outcomes, and how these effects vary in the cross-section of clientele and risk. Section 6 presents the complexity-yield puzzle and explores mechanism that potentially explain this puzzle. Section 7 concludes.

2 Institutional Background

The municipal bond market is distinct from other fixed income markets on both the supply side and the demand side. On the supply side, the market is heavily fragmented with over one million outstanding bonds issued by approximately 50,000 state and local governments. Bonds issuances finance a variety of projects that may be supported by the local tax base (general obligation bonds) or the cash flows generated by the underlying project (revenue

[5] Other recent work has found that municipal bond yields and trading outcomes are affected by the timeliness of trade reports (Chalmers et al., 2021), opioid abuse in the local community (Cornaggia et al., 2022b), the age of the local tax base (Butler and Yi, 2022), sea level rise exposure (Painter, 2020; Goldsmith-Pinkham et al., 2023), financial advisor conflicts of interest (Garrett, 2021), newspaper closures (Gao et al., 2020), and creditor protections for distressed municipalities (Gao et al., 2019).

[6] Related to complex design, Gao et al. (2022) find naive investors (proxied by low wealth) experience worse return outcomes relative to sophisticated investors when investing in complex structured products in the Chinese market.

bonds). On the demand side, retail investors comprise the single largest category of municipal bond investors, holding approximately 45% of outstanding bonds by par value. This is especially large among fixed income asset classes; for comparison, households hold only 3% of U.S. Treasuries and 1% of corporate bonds (Bagley et al., 2023c). Part of the reason for the high percentage of retail investors is that most states exempt municipal bond interest income from state taxes if the municipal bond was issued in that state, thereby motivating local investors to purchase in-state municipal bonds.

Retail investors are active on both the primary and secondary markets. On the primary market, retail investors are often encouraged to submit bids within a priority purchase window if the sale is negotiated through a single underwriter. Typical retail purchases take place through conventional third-party channels such as banks or brokers. These third parties may present investors with a curated list of higher-quality bonds to meet particular investment objectives; however, investors are encouraged to research the underlying features and risks of the bond via the official statements. For example, Edward Jones provides on their website a select list of municipal bond offerings that contains the names of the issuers and their official statements from the Municipal Securities Rulemaking Board (MSRB) (Edward Jones, 2023). For lower-quality bonds, the onus falls further on retail investors to conduct their own research by referring to the official statements directly through the MSRB. Immediately after a bond is issued, there is typically additional demand from retail investors to purchase bonds on the secondary market from dealers or mutual funds that act as intermediaries between the primary and secondary markets (Bagley et al., 2023b; Azarmsa, 2022).

Institutional investors are also highly active in the municipal bond market and hold the remaining 45% of bonds outstanding during our sample period. Institutional investors face similar challenges to retail investors when confronting the highly fragmented municipal bond market, processing disclosures, and choosing bonds that align with their investment objectives. Unlike most retail investors, however, institutional investors can achieve economies of scale in information processing through technology. For example, data vendors such as S&P Global and Refinitiv offer updated model-generated pricing estimates for illiquid municipal bonds to their clients. The cost of accessing this technology likely limits its use to institutional investors. In addition, many trading decisions require advanced information technology to identify trades in a timely manner. Buyers must be able to sift through large volumes of textual and numerical data and perform valuations across multiple dimensions, a process that requires a high degree of automation. From this standpoint, institutional investors are likely to approach the municipal bond market with a more sophisticated set of tools that allow them to make use of large quantities of data within a short period of time.

3 Theory and Empirical Predictions

In this section, we present a theoretical model that builds upon the canonical rational expectations model in Grossman and Stiglitz (1980) by introducing an exogenous disclosure requirement for a risky municipal bond issuer. In practice, a regulatory agency such as the SEC or MSRB may require issuing municipalities to provide more information in their official statements, thereby increasing the amount of information in but also the length and complexity of these statements. Sophisticated traders receive an imperfect signal of the final payoff of the risky municipal bond, and the disclosure requirement increases the precision of their signal. That is, sophisticated traders are able to process the additional disclosures to obtain a better forecast of the final payoff. Unsophisticated traders also receive an imperfect signal of the final payoff. However, the disclosure requirement reduces the precision of their signal because they have limited resources to process a large number of disclosures, and are unsure about which disclosures are most relevant to the final payoff. Put another way, unsophisticated investors are less willing to read official statements when they are faced with a document that is upward of 200 pages. Municipal bond markets are especially relevant to this setting because they are typically populated by more smaller, local investors that may not have the resources to process large, complex statements.[7]

Consider a setting with a risky municipal bond and a risk-free Treasury bond. Both assets are traded at $t = 1$, and the payoffs of these assets are realized at $t = 2$. The risk-free Treasury bond has a guaranteed payoff of 1 at $t = 2$, and is also available in unlimited supply. The risky municipal bond has an uncertain payoff of v which is realized at $t = 2$. We also assume that v is normally distributed with a mean of 0 and a precision of $\tau_v > 0$,

[7] One reason that municipal bond markets are populated by more smaller, local investors is that municipal bond interest income is exempt from state income taxes if the associated bond was issued in that state (Schultz, 2013; Babina et al., 2021).

where precision is the reciprocal of variance. In other words, we have that $v \sim N(0, \tau_v^{-1})$. The risky municipal bond is traded at $t = 1$ at an endogenous price p, and has a fixed supply $Q \geq 0$.

The municipal bond market is populated by three types of traders: sophisticated traders (I), unsophisticated traders (J), and liquidity traders. The total mass of the first two trader types is 1, the fraction of I traders is $\mu \in [0, 1]$, and the fraction of J traders is $1 - \mu$. The first two trader types have CARA utility for their terminal wealth at $t = 2$, and have a common risk aversion coefficient $\gamma > 0$. At $t = 0$, each sophisticated trader $i \in I$, our first trader type, receives a signal s_i about the terminal payoff v of the risky municipal bond that equals:

$$s_i = v + \epsilon_i, \text{ where } \epsilon_i \sim N(0, \delta^{-1}\tau_I^{-1}) \text{ and } \tau_I, \delta > 0. \tag{1}$$

Similarly, at $t = 0$, each unsophisticated trader $j \in J$, our second trader type, receives a signal s_j about the terminal payoff v of the risky municipal bond that equals:

$$s_j = v + u_j, \text{ where } u_j \sim N(0, (1-\delta)^{-1}\tau_U^{-1}) \text{ and } \tau_I \geq \tau_U > 0. \tag{2}$$

Importantly, the δ parameter measures the level of disclosure that is required for municipal bond official statements. By assumption, the precision of the signal for each trader $i \in I$ increases as δ increases, while the precision of the signal for each trader $j \in J$ decreases as δ increases. Lastly, the liquidity traders, our third trader type, demand x aggregate units of the risky municipal bond at $t = 1$, where $x \sim N(0, \tau_x^{-1})$ and $\tau_x > 0$. The liquidity traders do not pay any attention to the official statements, and their demand is uncorrelated with the true payoff v and the signals of the sophisticated traders and unsophisticated traders.

At $t = 1$, each non-liquidity trader chooses to invest in a quantity x of the risky municipality bond in order to maximize their expected utility conditional on the information that they receive from their respective signal (s_i, s_j) and the price of the risky municipal bond. The price of the municipal bond at $t = 1$ is determined by the market clearing condition that total demand from the sophisticated traders, unsophisticated traders, and liquidity traders equals the fixed supply of the municipal bond (Q). Each non-liquidity trader takes the optimizing behaviors of the other agents and the distributions of the random variables as given.

We conjecture that the price function at $t = 1$ takes the following linear form:

$$p = p_0 + p_v v + p_x x, \tag{3}$$

where p_0, p_v, and p_x will be determined by the equilibrium solution to the model.

First, we solve for the optimal demand of each sophisticated trader $i \in I$. Given the CARA-normal setup of our model, each sophisticated trader equivalently has mean-variance utility over their final wealth, and thus their demand D for the risky municipal bond takes the following form:

$$D_i(s_i, p) = \frac{E(v|s_i, p) - p}{\gamma Var(v|s_i, p)}. \tag{4}$$

For sophisticated traders and unsophisticated traders, observing p is equivalent to observing $s_p = (p - p_0)/p_v = v + (p_x/p_v)x = v + \rho^{-1}x$, where $\rho \equiv p_v/p_x$. Using Bayes' rule and the linear projection theorem, we find that the optimal demand for each sophisticated trader is:

$$D_i(s_i, p) = [\tau_I \delta s_i + \tau_x \rho^2 s_p - (\tau_I \delta + \tau_x \rho^2 + \tau_v)p]\gamma^{-1}. \tag{5}$$

Similarly, we find that the optimal demand for each unsophisticated trader is:

$$D_j(s_j, p) = [\tau_J(1 - \delta)s_j + \tau_x \rho^2 s_p - (\tau_J(1 - \delta) + \tau_x \rho^2 + \tau_v)p]\gamma^{-1}. \tag{6}$$

We obtain the equilibrium price p by substituting the above expressions for $D_i(s_i, p)$ and $D_j(s_j, p)$ into the market-clearing condition:

$$\int_0^\mu D_i(s_i, p)d_i + \int_\mu^1 D_j(s_j, p)d_j + x = Q. \tag{7}$$

We obtain an expression for p using this market clearing condition, match coefficients with equation (3) to obtain three equations that contain p_v, p_x, and p_0, and then solve this system of three equations to obtain closed-form expressions for p_v, p_x, and p_0. As a result, we obtain the following equilibrium price solution:

$$
\begin{aligned}
p &= p_v v + p_x x + p_0, \text{ where} \\
p_v &= B(\tau_x B + \gamma^2) \times D^{-1} \\
p_x &= \gamma(\tau_x B + \gamma^2) \times D^{-1}, \\
p_0 &= -Q\gamma \times D^{-1}, \\
D &= B(\tau_x B + \gamma^2) + \tau_v \gamma^2, \text{ and} \\
B &= \delta \times (\tau_I \mu - \tau_J(1-\mu)) + \tau_J(1-\mu).
\end{aligned} \tag{8}
$$

Lastly, we note that $B > 0$, which implies that D, p_v, p_x are also greater than zero, while p_0 is less than zero. Intuitively, this means that the price p is increasing in fundamental value v and noise trader demand x, and decreasing in bond supply Q. Given the parameters specified in our model, we have a unique and partially revealing noisy rational expectations equilibrium.

Most importantly, we can now determine how an exogenous increase in disclosure complexity δ affects equilibrium market outcomes. In this study, we focus on three outcomes. The first outcome is the cost of capital for the risky municipal bond (C), calculated as $E(v - p)$ (Easley and O'Hara, 2004; Hughes et al., 2007; Lambert et al., 2007). This is arguably our most important market outcome because a high cost of municipal debt can directly impact local infrastructure growth and local government budget constraints. Therefore, much of this study will focus on the effect of disclosure complexity on the cost of municipal debt. Our second outcome is market liquidity for the risky municipal bond (L), calculated as $1/p_x$. Intuitively, this measure captures the extent to which sophisticated and unsophisticated traders are willing to lean against orders from liquidity traders. If sophisticated and unsophisticated traders are less certain about true asset value on average, then they are less willing to absorb liquidity demand, indicating a less liquid market. Lastly, our third outcome is volatility of the risky municipal bond (σ), calculated as $(Var(v - p))^{1/2}$. Intuitively, if sophisticated and unsophisticated traders have more precise information on average, then the price of the risky municipal bond should deviate less from its fundamental value.

Calculating the above three metrics and then taking their derivatives with respect to disclosure complexity δ, we obtain the following results:

$$\begin{aligned}
\partial C/\partial\delta &= -\frac{C^2 \times (2\tau_x B + \gamma^2)}{Q\gamma} \times (\tau_I\mu - \tau_J(1-\mu)) && (9)\\
\partial L/\partial\delta &= \frac{L}{\gamma D} \times \frac{(\tau_x B + \gamma^2)^2 - \tau_x\tau_v\gamma^2}{\tau_x B + \gamma^2} \times (\tau_I\mu - \tau_J(1-\mu)) && (10)\\
\partial\sigma/\partial\delta &= -\frac{\gamma^2(D\tau_x^2 B + (2\tau_x B + \gamma^2)K)}{D^3\sigma\tau_x\tau_v} \times (\tau_I\mu - \tau_J(1-\mu)), \text{ where} && (11)\\
K &= \tau_v(\tau_x\tau_v\gamma^2 + (\tau_x B + \gamma^2)^2) - D\tau_x > 0.
\end{aligned}$$

In all three cases, notice that the sign of each derivative depends on whether or not $\tau_I\mu > \tau_J(1-\mu)$. The importance of this inequality makes economic sense because the $\tau_I\mu$ term represents the marginal value of disclosure complexity to sophisticated traders, while the $\tau_J(1-\mu)$ term represents the marginal cost of disclosure complexity to unsophisticated traders. Furthermore, for the cost of capital and volatility derivatives, all other variables in these equations are positive, meaning that the effects of complexity on the cost of capital and volatility depend only on the sign of $\tau_I\mu - \tau_J(1-\mu)$. We should also note that for the liquidity derivative, the second fraction can also be positive or negative depending on whether or not the τ_v term is extremely high relative to τ_x and γ. Because this theory is presented in the context of a risky municipal bond with an uncertain terminal payoff, we will assume that τ_v is not extremely high relative to τ_x and γ, meaning that the second fraction in the liquidity derivative is always positive. Therefore, under this mild assumption, the sign of the liquidity derivative also depends only on the sign of $\tau_I\mu - \tau_J(1-\mu)$. To summarize, we generate the following empirical predictions based on these derivatives:

Prediction 1(a): If $\tau_I\mu > \tau_J(1-\mu)$, then an increase in disclosure complexity will (i) reduce the cost of municipal debt, (ii) increase municipal bond liquidity, and (iii) reduce municipal bond price volatility.

Prediction 1(b): If $\tau_I\mu < \tau_J(1-\mu)$, then an increase in disclosure complexity will (i) increase the cost of municipal debt, (ii) decrease municipal bond liquidity, and (iii) increase municipal bond price volatility.

Prediction 1(c): If $\tau_I \mu = \tau_J(1-\mu)$, then an increase in disclosure complexity will not affect the cost of municipal debt, municipal bond liquidity, or municipal bond price volatility.

The remainder of this paper will focus on empirically testing the effect of disclosure complexity on market outcomes, and how these effects vary in the cross-section of bonds where there is variation in signal precision and the proportion of sophisticated traders in the market.

4 Data and Summary Statistics

4.1 Official Statements and Official Statement Complexity

Each municipal bond issue is accompanied by an official statement which is intended to be the primary source of information for investors to assess the potential risks of the investment. Through the official statements, issuers are legally required to disclose all relevant details on a bond issue so that investors in theory have complete information on that issue solely by reading the official statement (Feyer et al., 2018). The official statement can be thought of as similar to a prospectus for a firm planning to go public in that it provides a description of the associated securities (e.g., the amount offered and the bond maturity schedule), a means of marketing the investment, and all material information pertinent to the bonds such as the undelrying risks.

The official statement is required to be distributed to investors prior to the settlement of a transaction, per MSRB Rule G-32. While the official statement may be written with help from the underwriter, bond counsel, or municipal financial advisor, the legal liability associated with the official statement falls solely on the municipal issuer. An update to the SEC's Rule 15c2-12 requires all municipalities to post their official statements to the MSRB's Electronic Municipal Market Access (EMMA) website starting in July 2009. Therefore, we are able to obtain all municipal bond official statements through the EMMA dashboard from mid-2009 through October 2020. In order to access the text within the statements, we convert each document from PDF format into a machine readable format using the PDFTOHTML software program.[8] This process results in text data for 12,416,759 pages across 87,795

[8]This program can be accessed at https://pdftohtml.sourceforge.net/

official statements. The metadata included with the EMMA dashboard provide the 9-digit CUSIPs for each issue, allowing us to merge the data with our bond market characteristics described in Section 4.2.

The central text-related variable used in this study is complexity of the official statement. Given that we are among the first studies to conduct a textual analysis of municipal bond official statements, we do not take a firm stand on the dominant component(s) of complexity. Instead, we create an equal-weighted index of a variety of textual complexity measures: the total number of pages in the official statement, the total number of words, and several commonly used readability measures that provide the grade level required to understand the document, including the Flesch-Kincaid, Gunning-Fog, Automated Readability, Coleman Liau, Linsear Write, and SMOG grade level indices.[9] While these measures vary in their construction, they are highly correlated and tend to agree in the rankings of complexity between documents. To construct the complexity index, we first calculate the z-score for each measure so that is has a mean of zero and a standard deviation of one. The complexity index, which we denote *Complexity* in our tests, is constructed as the simple average of each of these z-scores which is then standardized again to have a mean of zero and a standard deviation of one for ease of interpretation in our later regression tests.

Figure 1 presents text samples from two official statements. Panel A provides a text sample from a below-median complexity statement from Alabama, while Panel B provides a text sample from an above-median complexity statement from Massachusetts. Both source documents require significant reader attention: the official statement from Alabama contains over 52,000 words across 144 pages and reads at an approximate 12th grade level, while the more complex document from Massachusetts contains almost 100,000 words and reads at an estimated 24th grade level. The latter document requires greater effort to interpret: the first sentence alone contains over 150 words with multiple parentheticals. From this standpoint, the information presented in the sample from Massachusetts is more complex than the sample from Alabama.

Table 1, Panel A presents summary statistics for our complexity index inputs. The

[9]The Flesch-Kincaid grade level index, for example, is calculated as $0.39 \times (\text{total words}/\text{total sentences}) + 11.8 \times (\text{total syllables}/\text{total words}) - 15.59$. See Table 2 of Wang et al. (2022) for the exact formulas and linguistic features associated with each of these measures.

average document reads at an approximately 21st grade level, which would be equivalent to a graduate degree. For context, a typical news article for the general public is written at a 12th grade level (Wasike, 2016). We also find that the average official statement in our sample contains 135 pages and 63,000 words. This high volume of information, in combination with the high average grade level needed to understand the document, suggests that a prospective investor would likely require multiple sittings to read and fully comprehend the official statement. This is potentially problematic for individual investors who want to "comparison shop" municipal bond offerings but lack the resources to fully process and integrate the information contained in multiple official statements.

4.2 Municipal Bond Primary and Secondary Market Data

We use the Mergent Municipal Bond Securities Database to identify primary market characteristics of municipal bonds. The bond-level data from Mergent include the offering yield, bond size, the issuance series, whether the bond was issued through the negotiated or competitive market, the issuance and maturity dates, credit ratings from Moody's and S&P (if the bond is rated), callable status, funding source (general obligation versus revenue), and whether the bond is insured. Because official statements cover several bonds within an issue, we aggregate these bond attributes to the issue level by calculating size-weighted averages. We supplement the bond data with census data on county-level population levels and unemployment rates as these local economic variables can also influence our outcome variables. We also supplement the bond data with secondary market trading information from the EMMA MSRB database, which contains the universe of over-the-counter transactions in municipal bonds. We use this secondary information to construct liquidity and volatility measures that we examine later in this study.

Table 1, Panel B presents summary statistics for the bond issue characteristics. Our merged sample starts in 2009, the first year of mandatory official statement postings to the EMMA platform by the MSRB, and ends in 2019. The average issue size is $32 million and the average maturity is 8.3 years. The average yield spread is 35 basis points, with an interquartile range that spans -1 basis point to 60 basis points.[10] A negative yield spread

[10]The issue-level offering yield spread is calculated as follows. For each municipal bond, we first calculate

is not uncommon in the municipal bond market because of the associated tax privileges, particularly in jurisdictions that exempt interest income from both federal and state income taxes. General obligation (GO) bonds, which are supported by the tax base of the issuer, represent 61% of issues in the sample, and the remaining 39% of issues are revenue bonds, which are backed by the cash flows generated by the underlying project. Under certain circumstances, issuers will purchase insurance to guarantee repayment of the bond: 13% of issues in our sample have insurance. Approximately half of the issues in our sample are bank qualified, which encourages investment banks to hold the bonds due to tax exemptions. Finally, about one-third of the issues in our sample are sold via negotiated sale: this involves a single underwriter who assembles an order book and determines the specific pricing of the bond. This method also involves a priority retail order period, which facilitates retail purchases. The remaining two-thirds of issues involve competitive sales in which multiple underwriters bid on the issue.

5 Baseline Empirical Results

The purpose of our empirical study is to test the effect of complexity on local government borrowing costs and other market outcomes. The central dependent variable used in this study is the issue-level offering yield spread (y), and the central independent variable used in this study is *Complexity*, a composite measure of textual complexity that is broadly based on two characteristics that are well-known to make a document costly to interpret: the total volume of information and the readability of the document (Loughran and McDonald, 2014a; Guay et al., 2016). Details on the construction of the offering yield spread and the complexity measure are provided in Section 4. According to our theoretical model, the effect of complexity on yield spread is unclear because it depends on whether or not the total

its yield spread as the difference between its offering yield and the coupon-equivalent U.S. Treasury yield. The coupon-equivalent US Treasury yield is calculated by taking the present value of the future bond payments using the zero-coupon U.S. Treasury yield curve from Gürkaynak et al. (2007) to obtain the price of the coupon-equivalent risk-free bond, and then obtaining the yield-to-maturity using this price and the bond's payment schedule (Longstaff et al., 2005). Lastly, we calculate the size-weighted average offering yield spread across bonds within each issue to obtain the municipal issue offering yield spread. For brevity, we will refer to this variable as the yield spread throughout the paper unless otherwise noted.

marginal value of complexity to the sophisticated investors exceeds the total marginal cost of complexity to the unsophisticated investors.

Formally, we test the following baseline ordinary least squares (OLS) regression model:

$$y_{it} = \beta \cdot Complexity_{it} + \delta_{sy} + \gamma \cdot X_{it} + \varepsilon_{it}, \tag{12}$$

where i is the bond issue, t is the year-month of issuance, δ_{sy} is a vector of state-year fixed effects, and X is a vector of control variables that includes the natural log of the total issue size, the natural log of the size-weighted years to maturity, indicator variables for whether the bond is general obligation (GO), insured, callable, and issued in the negotiated market (as opposed to the competitive market), the natural log of the population size in the county of issuance, and the unemployment rate in the county of issuance. Throughout our specifications, standard errors are double-clustered by county of issuance and year-month.

The results of our baseline tests are reported in Table 2. In column (1), we test the above regression model without controls or fixed effects, and we progressively add the control vector X in column (2), state fixed effects in column (3), and state-year fixed effects in column (4). Across all specifications, we find that there is a positive and statistically significant relationship between complexity and yield spread. Our most stringent specification in column (4) indicates that a one standard deviation increase in complexity is associated with a 4.9 basis point increase in the yield spread. In relative terms, this 4.9 basis point effect represents 7.9% of one standard deviation in the yield spread (62.3 basis points), or 5.1% of the average credit spread between AAA and BBB municipal bonds (95.5 basis points) during our sample period. For a typical issue with an issuance size of \$32 million and a duration of eight years, the 4.9 basis point effect represents an equivalent 39.2 basis point price effect, or about \$125,000 in additional present value interest payments. In aggregate terms, the 39.2 basis point price effect represents about \$1.4 billion in additional present value interest payments for our subsample of bonds that were issued during our sample period and have a complexity score of at least 0.50. Overall, our baseline results indicate that complex official statements are associated with higher municipal borrowing costs that are economically significant for local municipalities.[11]

[11] The control variable coefficients are largely expected and reflect findings in the prior literature. For

For illustrative purposes, we use the 39.2 basis point complexity-price effect in conjunction with our theoretical model to back out an estimate of the proportion of sophisticated traders in the municipal bond market (μ). We can then compare our structural estimate of μ to aggregate estimates provided by industry groups such as the Securities Industry and Financial Markets Association (SIFMA) as a reality check to the model. Methodologically, consider the cost of capital derivative in equation (9) with the following input parameters: normalized bond supply $Q = 1$, $\delta = 0.5$, $\gamma = 1$, $\tau_x = 10$, and $\tau_v = \tau_I = \tau_J = 0.33$. $\tau_x = 10$ is equivalent to one standard deviation in liquidity demand and calculated as $(1/10)^{0.5} = 0.32$ (relative to the normalized supply $Q = 1$). $\tau_v = \tau_I = \tau_J = 0.33$ is equivalent to one standard deviation in municipal bond price and calculated as $(1/0.33)^{0.5} = \$1.73$, which is approximately on par with one standard deviation in municipal bond price in the data. Lastly, $\delta = 0.5$ implicitly assumes that the sophisticated traders and unsophisticated traders have the same signal volatility before complexity is taken into account through the derivative $\partial C/\partial \delta$.[12]

Figure 2, Panel A provides a graph of our model-implied complexity-price effect ($\partial C/\partial \delta \times 0.35$) versus the proportion of sophisticated traders (μ). First, we find that the complexity-price effect is positive as long as $\mu < 0.5$, and that the positive complexity-price effect gets larger as μ decreases. Intuitively, the complexity-price effect gets larger as μ decreases because of the increasing total marginal cost of complexity to unsophisticated traders. As an example, if $\mu = 5\%$, then our model indicates that a one standard deviation increase in complexity would increase the cost of capital for municipal borrowers by 76 basis points. At the other extreme, if $\mu = 95\%$, then a one standard deviation increase in complexity would decrease the cost of capital for municipal borrowers by 76 basis points.

Importantly, when $\mu = 27\%$, the corresponding complexity-price effect of 39.2 basis points in Panel A of Figure 2 is the same as the empirical point estimate from our main regression specification in column (4) of Table 2. In reality, is this estimate of μ consistent with

example, the positive coefficient on the *Insured* indicator variable indicates that investors now require a premium for the average municipal bond during our post-financial crisis period, which is consistent with findings in Bergstresser et al. (2013) and Cornaggia et al. (2022a).

[12]For normalization purposes, we multiply this derivative by 0.35 because a one standard deviation increase in complexity corresponds to a 70% increase in the number of pages in the municipal bond official statement, or a 0.35 unit increase in the complexity variable δ relative to the baseline complexity δ of 0.50.

aggregate clientele statistics in the municipal bond market? According to SIFMA, individuals held about 47% of municipal bonds during our sample period, and the remaining 53% was held by mutual funds and ETFs (24%), banks (12%), insurance companies (13%), and an "other" group that includes state retirement funds and government-sponsored enterprises (4%). We assume that individuals are unsophisticated relative to institutions. We also assume that two-thirds of the mutual funds and ETFs are actively-managed and therefore sophisticated during our sample period (WSJ, 2019), and that banks, insurance companies, investors from the "other" group, and the remaining one-third of mutual funds and ETFs are passive liquidity investors. Therefore, the clientele breakdown in the municipal bond market is 47% unsophisticated investors, 24% $\times$ (2/3) = 16% sophisticated investors, and 37% liquidity traders. Therefore, the SIFMA statistics imply that $\mu = 16/(16+47) = 25.4\%$, which is reasonably close to our model implied estimate of 27%.

5.1 Complexity Effects in the Cross-Section of Risk

We next explore how the complexity-price effect varies in the cross-section of municipal bond coverage and risk. Official statement complexity is less likely to matter if external organizations such as credit rating agencies can process the relevant information from official statements and effectively communicate it to unsophisticated investors. If credit rating agencies can fully convey the relevant risks in easy-to-read credit rating reports, then those reports essentially act as perfect complexity-free substitutes for official statements. However, for risky municipal issues with opaque statements or underlying cash flows that are difficult to understand, credit rating agencies may not be able to fully convey the relevant risks in easy-to-read credit rating reports, and thus their reports may only act as partial substitutes for official statements. Therefore, we hypothesize that the complexity-price effect is stronger for unrated and lower-rated municipal issues.

Formally, we re-test the regression model in equation (12) using the following credit rating subsamples: (1) high-rated bonds, which are bonds in the top-two credit rating categories, (2) medium-rated bonds, which are bonds in the next two credit rating categories, (3) low-rated bonds, which are bonds in the remaining credit rating categories, and (4) unrated bonds, which are not rated by any credit rating agency. The results for each subsample test are reported in Table 3. In columns (1) and (2), we find that the complexity-yield effects for the high-quality and medium-quality groups are positive but very low at 1.3 basis points and 0.6 basis points, respectively.[13] By contrast, in column (3), we find a much larger complexity-yield effect of 7.4 basis points for low-rated bonds, indicating that credit rating reports do not provide a perfect substitute for risky municipal bonds with complex official statements, as those bonds have underlying cash flows that are more difficult to understand. Lastly, in column (4), we find a similarly large complexity-yield effect of 9.1 basis points for unrated bonds, indicating that the complexity-yield effect is especially strong when investors do not have access to easy-to-read credit rating reports.

Lastly, in Figure 2, Panel B, we graph the structural complexity-price effect versus precision in the fundamental value of the security (τ_v) while holding μ fixed at 27%. We continue to assume that the remaining parameter inputs are the same, and that $\tau_I = \tau_J = \tau_v$. This exercise allows us to examine how the complexity-price effect varies with the riskiness of the bond, which is inversely captured by the τ_I parameter on the x-axis. First, we find that the complexity-price effect is decreasing in τ_I, indicating that the effect of complexity on price is stronger for riskier bonds which have underlying cash flows that are harder to understand. Second, we can use our point estimates from Table 3 to back out estimates of τ_I. The complexity-yield effects documented in this table are equivalent to complexity-price effects of 7.0 basis points for high/medium-quality bonds, 59.2 basis points for low-quality bonds, and 72.8 basis points for unrated bonds. Using our graph in Figure 2, Panel B, we find that the implied structural estimates of τ_I for high/medium-quality, low-quality, and unrated bonds are 1.38, 0.21, and 0.16, respectively. That is, compared to high/medium-rated bonds, signal precision for low-rated bonds is about 85% lower, while signal precision for unrated bonds is about 88% lower. These results underline the importance of high-quality information disclosure for reducing harmful complexity-price effects for low-quality and unrated bonds.

We also compare the effects of complexity on yield for general obligation (GO) bonds and revenue bonds. General obligation bonds are backed by cash flows from the tax base of the issuing municipality. Investors have additional information about these cash flows that is complementary to the information in the official statement, and this is especially helpful

[13] In a pooled sample test, we find that the high-quality effect is not statistically significant from the medium-quality effect.

when the official statement is complex or difficult to process. Additional information that is relevant to the GO bond market includes publicly-available historical financial records on county government websites, tax records, intergovernmental revenue transfers, and employment statistics from the Bureau of Labor Statistics. By contrast, revenue bonds are only backed by the cash flows generated by a new project which lacks historical records or other information that would otherwise be helpful for interpreting complex official statements. We hypothesize that the complexity effect on yield will be more pronounced for revenue bonds, especially those that are low-rated or unrated and thus do not have reliable information from credit rating agencies.

We re-test the regression model in equation (12) using the following subsamples: (1) GO bonds, (2) high/medium-rated revenue bonds, (3) low-rated revenue bonds, and (4) unrated revenue bonds. The results for each subsample test are reported in Table 4. In column (1), we find that the complexity-yield effect for GO bonds is -2.4 basis points, indicating that the total marginal value of complexity for sophisticated traders is slightly greater than the total marginal cost for unsophisticated traders. This result suggests that sophisticated traders are more active in the GO bond market compared to the revenue bond market. In column (2), we find that the complexity-yield effect for high-quality revenue bonds is 3.7 basis points, indicating a higher marginal cost of complexity for unsophisticated traders when there is no supporting public information outside what is supplied by credit rating agencies. Lastly, in columns (3) and (4), we find large complexity-yield effects of 10.5 basis points and 14.0 basis points for low-quality revenue bonds and unrated revenue bonds, respectively. These results suggest that complexity is especially costly when unsophisticated investors do not have access to complementary information from historical financial records or credit rating reports and thus can only rely on complex official statements.

5.2 Complexity Effects in the Cross-Section of Clientele

The results in the previous subsection indicate that the marginal cost of complexity for unsophisticated traders exceeds the marginal value for sophisticated traders for most issues ($\tau_J(1-\mu) > \tau_I\mu$), especially those that lack supporting information from credit rating agencies or other public sources. In this section, we focus on the μ parameter in the cross-section of issues where there is variation in the proportion of sophisticated investors. Our model-implied estimates in Figure 2, Panel A indicate that the complexity-yield premium for municipal debt is increasing in the proportion of unsophisticated investors $(1-\mu)$. The purpose of this section is to empirically test if this relationship holds in the data.

We start by exploiting cross-state variation in tax privileges provided to local municipal bond investors. A unique feature of the municipal bond market is that most states allow investors to exempt municipal bond interest income from their state income taxes as long as that bond was issued in their state of residence. This tax privilege is especially appealing for local individual investors who reside in states with high income tax rates. On the other hand, for states with zero income tax rates such as Texas or Washington, there is no tax privilege for holding municipal bonds, meaning that local individual investors are less likely to represent a large percentage of the municipal bond investor base. The same holds true for the four states that do not provide tax privileges for holding municipal debt (i.e., Illinois, Iowa, Oklahoma, and Wisconsin). Importantly, Babina et al. (2021) show that municipal bonds issued in the nine states in the lowest quintile of state tax privilege, which includes the six states mentioned above and also Alaska, Indiana, and Nevada, have significantly lower in-state ownership compared to the remaining four quintiles. Under the assumption that bonds issued in these "no tax privilege" states are less likely to be held by unsophisticated local investors, and thus more likely to be held by sophisticated institutional investors that can process complex official statements, we hypothesize that the complexity-yield premium is lower in these no-privilege states.

We re-test the regression model in equation (12) using the following subsamples: (1) the nine states in the lowest quintile of tax privilege from Babina et al. (2021), and (2) the states in the remaining four quintiles of tax privilege. The results are reported in columns (1) and (2) of Table 5. Column (1) indicates that the complexity-yield premium is positive (2.3 basis points) but statistically insignificant for states in the lowest quintile of tax privilege. By contrast, column (2) indicates that the complexity-yield premium is large (5.6 basis points) and highly statistically significant for states in the remaining quintiles of tax privilege. In column (3), we re-test the same regression model using our full sample and also interact complexity with an indicator variable for the no-privilege states, and confirm that the complexity-yield premium in no-privilege states is close to zero and also statistically significantly lower than the complexity-yield premium in the remaining tax-privilege states. Lastly, in column (4), we re-test the regression model in column (3), except that we use states that had a zero state tax rate (and therefore no tax privilege) in 1947 as our no-privilege states. This empirical strategy is used in Babina et al. (2021) to address the potential endogeneity concern that unobserved factors may be driving both current tax privilege policy and the municipal bond yield spread. Using tax privilege in 1947 addresses this endogeneity concern because modern tax privilege is likely a vestige of historical tax privilege, and historical tax privilege is unlikely to be correlated with modern yield spreads. In column (4), we find that the complexity-yield premium does not exist in states that had no tax privilege in 1947, and that the complexity-yield premium is positive (6.7 basis points) and statistically significant in the remaining states. Overall, the results in this table indicate that the complexity-yield premium is positive and significant in states that offer favorable tax treatment for holding local municipal bonds and thus attract relatively less sophisticated local investors.

We provide further evidence of the importance of clientele makeup for the complexity-yield premium by exploiting cross-sectional variation in bond issue type. First, we consider bank-qualified (BQ) municipal issues which are designated to municipalites that issue less than \$10 million in a calendar year. Normally, banks cannot deduct municipal interest income for tax purposes, but BQ municipal issues are an exception. The tax privilege provided by the BQ status increases the likelihood that sophisticated institutional investors will invest in those bonds, thereby increasing μ. Indeed, Dagostino (2022) shows that BQ issues just under the \$10M size threshold are significantly more likely to be held by banks compared to non-BQ issues just over the \$10 million size threshold. Therefore, we expect that BQ bonds will have a lower complexity-yield premium due to a greater proportion of sophisticated institutional investors. Second, we consider municipal issues with an average offering price of at least \$103 per bond. According to Bagley et al. (2023a) and Bond Buyer (2022), institutional investors prefer to purchase bonds with large premiums to avoid the "de minimis risk" that the municipal bond falls below a pre-specified price threshold which triggers ordinary income tax rates on price appreciation. Furthermore, some mutual funds even have informal protocols to sell municipal bonds that fall below \$103 due to de minimis risk. Therefore, we expect that sophisticated institutional investors are more likely to hold municipal issues if they have an offering price of at least \$103, which would imply a lower complexity-yield premium for these high-priced bonds. Lastly, we consider municipal bonds that are issued in the negotiated market as opposed to the competitive market. According to WM Financial Strategies, negotiated bonds typically have lower credit quality or unusual financing terms, and thus require special handling by an underwriter as opposed to being sold in the competitive market. Furthermore, according to MSRB, retail investors are also given a "priority purchase window" for negotiated bonds which gives them bidding priority over institutional investors in some circumstances. Therefore, we expect a higher complexity-yield premium for negotiated bonds due to a combination of more unsophisticated retail investors and riskier cash flows that are more difficult to process.

The results of these cross-sectional tests are reported in Table 6. In column (1), we re-test the regression model in equation (12) with the inclusion of a Complexity $\times$ BQ interaction term, where BQ is an indicator variable that equals one if the bond is bank-qualified. We also include the standalone BQ indicator variable in our set of controls. Column (1) indicates that there is a significant complexity-yield premium of 6.6 basis points for non-BQ bonds, and no complexity-yield premium for BQ bonds. In column (2), we replace the BQ indicator variable from column (1) with the High Price indicator variable and find evidence of a significant complexity-yield premium of 9.1 basis points for lower-priced bonds, and no complexity-yield premium for High Price bonds. In column (3), we replace the BQ indicator variable from column (1) with the Negotiated indicator variable and find evidence of a significant complexity yield premium of 7.1 basis points for non-negotiated (competitive) bonds, and an even larger complexity-yield premium of 14.1 basis points for negotiated bonds. Lastly, in column (4), we include all of the additional interaction terms and control variables from columns (1) to (3) and find qualitatively similar results, although a small complexity-yield premium still persists for BQ bonds in this test (2.2 basis points), even if it is significantly lower than the complexity-yield premium for non-BQ bonds (10.5 basis points). Overall, these results support our earlier tests on tax-induced clientele effects, and further highlight the importance of sophisticated institutional investors for reducing the complexity-yield premium.

5.3 Complexity Effects on Volatility and Liquidity

Our theoretical model also predicts that complexity increases yield volatility and reduces liquidity in the municipal bond market as long as the total marginal cost of complexity to unsophisticated traders exceeds the total marginal value to sophisticated traders. Given the evidence in the last section showing a positive complexity-yield premium, this latter condition appears to hold. Therefore, we predict that complexity will also increase yield volatility and reduce liquidity. In this section, we empirically test these predictions.

We start by examining the effect of complexity on yield volatility. To measure yield volatility for each bond, we calculate the standard deviation of the secondary yield spread using municipal bond transactions during the three-year period immediately after the bond is issued. Secondary yields are obtained from MSRB, and the secondary yield spread is calculated similarly to the offering yield spread by subtracting the coupon-equivalent risk-free rate. We obtain issue-level measures of yield volatility by calculating average yield volatility across all bonds within each issue. Across all issues in our sample, we find that issue-level yield volatility has a mean of 24.6 basis points and a standard deviation of 20.8 basis points.

We re-test the regression model in equation (12), except that we use issue-level yield volatility as our dependent variable. The results are reported in column (1) of Table 7. We find that a one standard deviation increase in complexity is associated with a 0.3 basis point increase in yield volatility, or 1.4% of one standard deviation in yield volatility. The directional effect is consistent with our theoretical prediction, although the magnitude of the effect is somewhat modest. One potential reason for the modest effect is that we lose about one-third of our observations due to a lack of secondary market trades in some issues. For example, we find that 51% of the lost observations are unrated issues, while only 11% of the remaining observations are unrated issues, implying that our reduced sample is biased toward rated issues which are less likely to be affected by complexity in the higher rating categories. Therefore, to account for the biased sample, we re-test the regression model in column (1), but also include Complexity × Unrated, Complexity × Mid Rating, and Complexity × Low Rating. The results are reported in column (2). First, we find near-zero complexity-volatility effects for high-rated and medium-rated bonds, which is consistent with our earlier results showing that the complexity-yield premium is also near-zero for these ratings classes. Second, we find that the complexity-volatility effects for low-rated and unrated bonds are larger at 0.8 basis points (3.8% of one standard deviation in yield volatility) and 2.0 basis points (9.6% of one standard deviation in yield volatility), respectively. In columns (3) and (4), we also test the effects of complexity on yield volatility for general obligation bonds versus revenue bonds and high-priced bonds versus low-priced bonds, respectively. Similar to our previous results on the complexity-yield premium, we find that complexity has positive and significant effects on yield volatility for revenue bonds and low-priced bonds, but no effects on yield volatility for GO bonds and high-priced bonds. Overall, our results indicate that complexity increases yield volatility, particularly if the issue lacks complementary information from credit rating agencies or other public sources, or if institutional investors are less likely to be active in the market for that issue.

Lastly, we test the effect of complexity on liquidity. Liquidity for each bond is calculated as the average roundtrip transaction cost in the first three years post-issuance, where roundtrip transaction cost is the percentage difference between a customer buy order and a same-size customer sell order within a short time window.[14] Issue-level liquidity is calculate as average liquidity across all bonds within that issue and denoted $Markup$. We also calculate issue-level liquidity for retail roundtrip transactions only ($RMarkup$), which are transactions below $100 thousand in trade size, and institutional roundtrip transactions only ($IMarkup$), which are transactions greater than or equal to $100 thousand in trade size. To maintain a balanced panel for our liquidity analysis, we require the issue to have positive, non-empty observations for these liquidity measures. We also exclude unrated bonds from this analysis because most do not have sufficient liquidity observations across the three categories. Across all issues in this analysis, we find that the mean (standard deviation) of $Markup$ is 102.9 basis points (56.1 basis points), the mean (standard deviation) of $RMarkup$ is 110.7 basis points (74.4 basis points), and the mean (standard deviation) of $IMarkup$ is 67.5 basis points (67.3 basis points). Consistent with the previous literature, we find that the average markup for retail trades is significantly higher than the average markup for

[14]We follow the methodology in Green et al. (2007) to construct this measure. We find a median roundtrip transaction cost of approximately 90 basis points during our sample period, which closely aligns with results reported in Griffin et al. (2023).

institutional trades.

We re-test the regression model in equation (12), except that we use $Markup$, $RMarkup$, and $IMarkup$ as our dependent variables. The results are reported in Table 8. In column (1), we find that complexity reduces liquidity, which is consistent with our theoretical prediction. In particular, we find that a one standard deviation increase in complexity is associated with a 1.3 basis point increase in $Markup$, or 2.3% of one Markup standard deviation. Similar to our volatility results, the directional effect is consistent with our theoretical prediction, although the effect is somewhat modest, likely due to the fact that conditioning our sample on observed liquidity biases the sample toward safer bonds that are more frequently-traded and less affected by complexity. In column (2), we find that complexity effect on $RMarkup$ is slightly stronger (1.7 basis points, or 2.3% of one standard deviation in $RMarkup$), while the complexity effect for $IMarkup$ is negative (-1.1 basis points, or -1.6% of one standard deviation in $IMarkup$). In the context of our theoretical model, these opposite-direction results make sense because sophisticated institutional traders are more willing to participate in the market and take the other side of liquidity-motivated trades when complexity is higher, while unsophisticated retail traders are less willing to participate for the same reason. Therefore, the markup differential between retail and institutional traders is wider for issues with more complex official statements.

6 The Complexity-Yield Puzzle

Our evidence suggests that a large cross-section of municipal bond issuers could decrease their borrowing costs by simplifying their official statements, thereby catering to the large population of unsophisticated retail investors. However, despite these potential savings, complexity has been steadily increasing over time. In particular, Figure 3, Panel A indicates that complexity has increased by about 0.5 standard deviations from 2010 to 2019, while Panel B indicates that residual complexity (obtained from a regression of complexity on the bond-level control variables from our baseline regression specification) has increased in a similar fashion. In terms of raw inputs, we find that the average number of pages in the official statement has increased by about 45 (40%), the average number of words has increased by about 30 thousand (64%), and the average Gunning-Fog readability grade has increased by about 1.8 (9.0%).

Given our documented evidence of a positive and significant complexity-yield premium, why has complexity increased over time? In this section, we consider two hypotheses to answer this question. Our first hypothesis is that municipal bond issuers are increasingly catering their official statements to sophisticated institutional investors who extract more precise signals from complex official statements. Supporting this hypothesis is evidence from Bergstresser and Cohen (2016) of a declining trend in the percentage of households owning municipal debt, and an increasing trend in the percentage of municipal debt held by households in the upper 0.5% of the wealth distribution. These findings suggest that the percentage of sophisticated investors in the municipal bond market has been increasing over time, and our first hypothesis is based on the idea that municipalities are increasingly catering to sophisticated investors by releasing more detailed and complex official statements. Our second hypothesis is that regulation in the municipal bond market from the SEC, the MSRB, and state-level agencies has become increasingly burdensome over time, thereby imposing additional complexity in official statements. Supporting this hypothesis is evidence from Dyer et al. (2017) that the increasing length and decreasing readability of 10-K filings for publicly-traded firms is mostly attributable to new SEC and FASB disclosure requirements.

Our first step is to examine the evolution of the complexity-yield premium over time. Under the catering hypothesis, if the proportion of sophisticated investors is increasing and municipalities are adjusting official statement complexity to cater to this set of investors, then we should observe a decrease in the complexity-yield premium in the time series. We re-test the regression model in equation (12), except that we include interactions of complexity and each full year in our sample. The results in Figure 4 indicate that the complexity-yield premium has remained fairly steady over time at approximately 5.0 basis points, with no statistically significant changes over time. Therefore, this evidence suggests that issuers have not catered to the increasing percentage of sophisticated investors to extract a lower yield.

Our second step is to examine the evolution of complexity across our four credit rating categories. Under the regulation hypothesis, complexity trends should not differ across credit rating categories because regulatory changes implemented by the SEC or MSRB are applicable to all issuers. In contrast, an increase in complexity for only the higher rated issuances would suggest municipalities are adding more information to bonds that do not carry a high complexity-yield premium. Consistent with the regulation hypothesis, we find in Figure 5 that the complexity trends are near-identical across credit rating categories, with no statistically significant differences in the slopes of any two trendlines. Of course, it is plausible that the upward trends in complexity are also consistent with the catering hypothesis if issuers from all credit rating categories cater to the growing sophisticated investor clientele. Therefore, we also examine trends in complexity for BQ versus non-BQ bonds and high-price versus lower-price bonds, with the reasoning that the growing sophisticated investor base is more likely to lean toward the BQ bonds and high-price bonds, implying a stronger trend for these bonds. In unreported tests, we also find that the trends are similar across all four of these categories, suggesting that clientele change is less likely to be the main driver of the rising complexity. Lastly, we examine differences in complexity trends for bonds located in no-tax privilege states versus the remaining states. In this case, we find that the upward trend in complexity is about 38% stronger for issuers located in no-tax privilege states compared to tax-privilege states. This evidence suggests that the clientele hypothesis is more likely to explain some of the upward trend in complexity in no-tax privilege states.

Our third step is to test the regulatory hypothesis in the cross-section of states. During the 2015 to 2017 period under SEC Chair Mary Jo White, there was a strong push to increase the number of enforcement officers in regional SEC offices (White, 2016). As a result, the number of enforcement officers increased by 70% on average.[15] However, there was also strong regional variation in enforcement officer growth as a result of this push, with the Boston office seeing an increase of 150% and the Forth Worth office seeing an increase of only 9%. If the regulatory hypothesis holds, then we would expect a significant increase in complexity after the regulatory push in early 2015 for states with regional offices that experienced high enforcement officer growth relative to states with regional offices that experienced low enforcement officer growth.

Methodologically, we define states as "high enforcement growth" (HEG) if they experienced above-median enforcement growth relative to other states.[16] We then regress com-

[15] We obtain the quarterly SEC employment trend by office location and job function from the Fedscope – Federal Workforce Data. This data enables us to track changes in the employment of SEC enforcement officers at the state level.

[16] HEG SEC offices are Boston, Chicago, Philadelphia, San Francisco, Los Angeles, and Salt Lake City.

plexity on indicator variables for each year, interactions of each year with an HEG indicator variable, and the control variables used in our main specification. The coefficient on each $Year \times HEG$ interaction term represents the difference in complexity for HEG states relative to non-HEG states in that year. Figure 6 provides a graph of the point estimates on these interaction terms versus year. Prior to the enforcement push, average complexity in HEG states is not significantly different from average complexity in non-HEG states. During the first year of the enforcement push in 2015, we also observe no significant differences in complexity. Given that the enforcement push was staggered over three years, and that many issuers may have released official statements early in that same year and prior to any enforcement push, this result is unsurprising. However, starting in 2016, we find that complexity in HEG states is about 0.15 standard deviations higher than complexity in non-HEG states, and that this effect persists to the end of our sample period. Therefore, our evidence suggests that regulatory enforcement is an important driver of complexity, thereby providing additional support for the regulatory hypothesis.

Lastly, we decompose complexity into multiple topics in order to test if the complexity-yield premium loads heavily on any particular topic. We use Latent Dirichlet Allocation (LDA) (Blei et al., 2003) in order to uncover topics discussed within the OS text. LDA is an unsupervised machine learning model that represents each individual text as a weighted collection of topics, where the topics are determined by the distribution of individual words across the corpus. Since the model uses only the corpus itself (i.e., samples of official statements), the construction of topics takes place without additional decisions regarding the subject matter: this removes the potential for biases that could arise if we were to create topics via a list of predetermined keywords, for example. While the model determines each topic's composition, we ultimately assign labels to the topics in order to provide interpretation.

The LDA approach is especially useful in the analysis of municipal bond official statements since the documents are not organized in a uniform manner. For instance, unlike 10-K filings, there is no specific risk disclosure section that can be consistently identified across statements. Since different sections contain different types of disclosures within the official statements, we would expect topic weightings to vary substantially within the document. We, therefore, estimate the LDA model at the OS-page level for the purpose of recovering topic weights for each page. Specifically, we first use LDA to group textual information into 100 labels for each OS-page.[17] We then assign each topic to four broad topic categories: (1) risk, (2) bond description, (3) cash flow, and (4) legal. In Table 9, we provide a sample list of topics identified by the LDA model, and we place each of these LDA topics under the four broad topic categories. As illustrative examples, we assign the LDA topics of liability, retirement, and pension to the "Risk" category, housing, improvement, and swap to the "Bond Description" category, municipality, mayor, and commissioner to the "Local" category, and agreement, contract, and pursuant to the "Legal" category. We then construct an OS-page level topic weighted complexity measure. In particular, for each OS-page, we compute the product of the complexity measure and topic proportion (e.g., if OS A, page 1 has risk topic proportion = 0.1, complexity index = 10, the risk topic weighted complexity measure is then 0.1 times 10). We then aggregate these measures to the OS level (CUSIP-8) and compute z-scores using all issues.

For each credit rating category, we re-test the regression model in equation (12), except that we replace our broad complexity measure with the four topical complexity measures described above. The results are reported in Table 10. In column (1), we find that only legal complexity explains yield for high-quality bonds, with a 4.5 basis point increase in yield for every standard deviation increase in legal complexity. In column (2), we find that legal complexity is the main driver of yield for medium-quality bonds (3.3 basis points), although description complexity also has a small effect. Lastly, in columns (3) and (4), we again find that legal complexity is the main driver of yield for low-quality and unrated bonds (12.3 basis points and 19.8 basis points, respectively), and that the economic magnitudes of the legal-complexity effect is much stronger compared to the higher-quality categories. For low-quality and unrated bonds, we also find that description complexity and risk complexity matter for yield. Overall, our results suggest that legal complexity is the most important driver of yield across all credit categories. Given that regulation is often connected to legal matters, our results further suggest that complexity is strongly connected to regulatory burden.

[17] A LDA label is the dot product between a vector of weights and a vector of words (e.g., topic X= $0.4 \cdot$ "*entity*" $+ 0.3 \cdot$ "*budget*" $+ 0.2 \cdot$ "*fund*" + ...). Each OS-page is made up of the proportions of each topic.

7 Conclusion

Effective information disclosure in financial markets is crucial for the efficient allocation of capital. In this study, we argue that issuers would incur lower borrowing costs if they tailor their disclosures to their clientele base. Unsophisticated retail investors may not have the resources to process long, complex official statements; if the clientele base leans heavily toward this investor type, then the issuer would be better off releasing shorter, simplified official statements in order to extract a lower cost of capital.

The municipal bond market is ideal for testing the effect of official statement complexity on the cost of capital because it is uniquely populated by a large number of unsophisticated retail investors. We textually process over 80,000 municipal bond official statements to determine the level of complexity in each of these statements, and provide evidence of an economically significant complexity-yield premium. Importantly, we find that the size of the complexity-yield premium very much depends on the clientele base. In particular, when sophisticated institutional investors comprise a larger percentage of the clientele base, which is the case for bonds that are bank-qualified or issued in states without any tax privileges for local investors, then we find evidence of a near-zero complexity-yield premium. Conversely, the complexity-yield premium is much larger when the clientele base exogenously leans toward more unsophisticated retail investors. We also find that the complexity-yield premium is exacerbated if the official statement lacks complementary information from credit rating agencies or other public sources, highlighting an additional challenge that lower-quality issuers face when attempting to raise capital for local infrastructure projects.

Despite the economically significant complexity-yield premium, we find that official statement complexity has been steadily rising during our sample period. We test two hypotheses to uncover the reasons for this puzzling trend: (1) issuers are increasingly catering to the growing percentage of sophisticated investors in the municipal bond market (Bergstresser and Cohen, 2016), and (2) increasing regulatory burden requires issuers to release increasingly complex statements (Dyer et al., 2017). Using cross-state variation in the SEC push for greater regional enforcement starting in 2015, we find that the post-2015 rise in complexity is stronger in states with greater enforcement growth, supporting the regulatory hypothesis. However, the catering hypothesis also appears to have some support for states that offer no tax privileges for local municipal bond investors. When we decompose complexity into multiple topics using an LDA generative statistical model, we find that legal complexity is the most important driver of the complexity-yield premium, further supporting the regulatory hypothesis.

Our findings suggest that additional costs are being imposed on municipal bond issuers, especially those that operate in markets with more unsophisticated retail investors or those that have less complementary information from external sources such as credit rating agencies. One possible solution for reducing the costs associated with the complexity-yield premium is to require issuers to provide a brief executive summary of the most relevant risks for potential lenders at the beginning of the official statement so that resource-constrained investors do not necessarily have to pore through long, complicated statements to identify the relevant risks. For smaller municipalities that are less likely to obtain credit ratings due to budgetary constraints, a streamlined executive summary that outlines key risks would be especially effective in reducing their borrowing costs and promoting economic growth.

References

Adelino, M., Cunha, I., Ferreira, M. A., 2017. The economic effects of public financing: Evidence from municipal bond ratings recalibration. The Review of Financial Studies 30, 3223–3268.

Azarmsa, E., 2022. Financing infrastructure with inattentive investors: The case of US municipal governments. Working paper. University of Oklahoma.

Babina, T., Jotikasthira, C., Lundblad, C., Ramadorai, T., 2021. Heterogeneous taxes and limited risk sharing: Evidence from municipal bonds. The Review of Financial Studies 34, 509–568.

Bae, J. W., Belo, F., Li, J., Lin, X., Zhao, X., 2022. The opposing effects of complexity and information content on uncertainty dynamics: Evidence from 10-K filings. Management Science, forthcoming.

Bagley, J., Gissler, S., Hiteshew, K., Ivanov, I. T., 2023a. Pushing bonds over the edge: Monetary policy and municipal bond liquidity. Working paper. Municipal Securities Rulemaking Board.

Bagley, J., Rosero, T. Q., Vieira, M., 2023b. Trading patterns in the primary vs. secondary market for municipal bonds. https://www.msrb.org/sites/default/files/2023-05/Trading-Patterns-in-the-Primary-vs-Secondary-Market-for-Municipal-Bonds.pdf.

Bagley, J., Vieira, M., Hamlin, T., 2023c. Trends in municipal bond ownership. https://www.msrb.org/sites/default/files/Trends-in-Municipal-Securities-Ownership.pdf.

Ben-Raphael, A., Da, Z., Israelsen, R. D., 2017. It depends on where you search: Institutional investor attention and underreaction to news. Review of Financial Studies 30, 3009–3047.

Bergstresser, D., 2023. The municipal bond market. In: *Research Handbook of Financial Markets*, Edward Elgar Publishing, pp. 301–330.

Bergstresser, D., Cohen, R., 2016. Changing patterns in household ownership of municipal debt: Evidence from the 1989-2013 surveys of consumer finances. Hutchins Center Working Paper.

Bergstresser, D., Cohen, R., Shenai, S., 2013. Demographic fractionalization and the municipal bond market. Municipal Finance Journal 34, 1–38.

Blankespoor, E., deHaan, E., Marinovic, I., 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. Journal of Accounting and Economics 70, 101344.

Blei, D. M., Ng, A. Y., Jordan, M. I., 2003. Latent dirichlet allocation. Journal of Machine Learning Research 3, 993–1022.

Bloomberg, 2022. Retail traders rush into bond markets in hunt for juicier yields. https://www.bloomberg.com/news/articles/2022-11-14/retail-traders-rush-into-bond-markets-in-hunt-for-juicier-yields.

Bodnaruk, A., Loughran, T., McDonald, B., 2015. Using 10-K text to gauge financial constraints. Journal of Financial and Quantitative Analysis 50, 623–646.

Boehmer, E., Jones, C. M., Zhang, X., Zhang, X., 2021. Tracking retail investor activity. The Journal of Finance 76, 2249–2305.

Bond Buyer, 2022. Potential disclosure issues with discount munis? https://www.bondbuyer.com/opinion/potential-disclosure-issues-with-discount-munis.

Brancaccio, G., Kang, K., 2022. Search frictions and product design in the municipal bond market. Working paper. National Bureau of Economic Research.

Brown, N. C., Crowley, R. M., Elliott, W. B., 2020. What are you saying? Using topic to detect financial misreporting. Journal of Accounting Research 58, 237–291.

Butler, A. W., Yi, H., 2022. Aging and public financing costs: Evidence from US municipal bond markets. Journal of Public Economics 211, 1–18.

Bybee, L., Kelly, B. T., Manela, A., Xiu, D., 2021. Business news and business cycles. Working paper. National Bureau of Economic Research.

Casey, A. J., Macey, J. C., 2020. The Hertz maneuver (and the limits of bankruptcy law). University of Chicago Law Review Online. https://www.wsj.com/articles/private-investors-buy-up-retail-real-estate-as-bigger-players-remain-cautious-11655208001.

Chalmers, J., Liu, Y., Wang, Z. J., 2021. The difference a day makes: Timely disclosure and trading efficiency in the muni market. Journal of Financial Economics 139, 313–335.

Cohen, L., Malloy, C., Nguyen, Q., 2020. Lazy prices. The Journal of Finance 75, 1371–1415.

Cornaggia, J., Cornaggia, K. J., Israelsen, R. D., 2018. Credit ratings and the cost of municipal financing. The Review of Financial Studies 31, 2038–2079.

Cornaggia, K., Hund, J., Nguyen, G., 2022a. The price of safety: The evolution of municipal bond insurance value. Management Science, forthcoming.

Cornaggia, K., Hund, J., Nguyen, G., Ye, Z., 2022b. Opioid crisis effects on municipal finance. The Review of Financial Studies 35, 2019–2066.

Cuny, C., 2018. When knowledge is power: Evidence from the municipal bond market. Journal of Accounting and Economics 65, 109–128.

Cuny, C., Li, K., Nakhmurina, A., Watts, E. M., 2022. The information content of municipal financial statements: Large-sample evidence. Working paper. New York University.

Dagostino, R., 2022. The impact of bank financing on municipalities' bond issuance and the real economy. Working paper. University of Rochester.

De Sousa, A., 2021. Retail trends in crypto. Federal Reserve Bank of San Francisco, Regional Banker Forum.

Dyer, T., Lang, M., Stice-Lawrence, L., 2017. The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. Journal of Accounting and Economics 64, 221–245.

Easley, D., O'Hara, M., 2004. Information and the cost of capital. The Journal of Finance 59, 1553–1583.

Edward Jones, 2023. Municipal bond offerings. https://www.edwardjones.com/us-en/investment-services/investment-products/fixed-income-investments/prospectuses-pricing/municipal-bond-offerings.

Ertugrul, M., Lei, J., Qiu, J., Wan, C., 2017. Annual report readability, tone ambiguity, and the cost of borrowing. Journal of Financial and Quantitative Analysis 52, 811–836.

Farrell, M., Green, T. C., Jame, R., Markov, S., 2022. The democratization of investment research and the informativeness of retail investor trading. Journal of Financial Economics 145, 616–641.

Feyer, R., Hernandez, J., Radecki, A., Spitz, S., 2018. Disclosure obligations of issuers of municipal securities. https://www.orrick.com/en/Insights/2018/10/Disclosure-Obligations-of-Issuers-of-Municipal-Securities.

Fishman, M. J., Hagerty, K. M., 2003. Mandatory versus voluntary disclosure in markets with informed and uninformed customers. Journal of Law, Economics, and Organization 19, 45–63.

Gao, P., Hu, A., Kelly, P., Peng, C., Zhu, N., 2022. Asset complexity and the return gap. Forthcoming, Review of Finance.

Gao, P., Lee, C., Murphy, D., 2019. Municipal borrowing costs and state policies for distressed municipalities. Journal of Financial Economics 132, 404–426.

Gao, P., Lee, C., Murphy, D., 2020. Financing dies in darkness? The impact of newspaper closures on public finance. Journal of Financial Economics 135, 445–467.

Garrett, D., 2021. Conflicts of interest in municipal bond advising and underwriting. Working paper. University of Pennsylvania.

Garrett, D., Ordin, A., Roberts, J. W., Suárez Serrato, J. C., 2023. Tax advantages and imperfect competition in auctions for municipal bonds. Review of Economic Studies 90, 815–851.

Goldsmith-Pinkham, P., Gustafson, M. T., Lewis, R. C., Schwert, M., 2023. Sea-level rise exposure and municipal bond yields. Review of Financial Studies, forthcoming.

Goldstein, I., Yang, L., 2017. Information disclosure in financial markets. Annual Review of Financial Economics 9, 101–125.

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

Griffin, J. M., Hirschey, N., Kruger, S., 2023. Do municipal bond dealers give their customers "fair and reasonable" pricing? The Journal of Finance 78, 887–934.

Grossman, S. J., Stiglitz, J. E., 1980. On the impossibility of informationally efficient markets. The American Economic Review 70, 393–408.

Guay, W., Samuels, D., Taylor, D., 2016. Guiding through the fog: Financial statement complexity and voluntary disclosure. Journal of Accounting and Economics 62, 234–269.

Gürkaynak, R. S., Sack, B., Wright, J. H., 2007. The US treasury yield curve: 1961 to the present. Journal of Monetary Economics 54, 2291–2304.

Hellwig, M., 1980. On the aggregation of information in competitive markets. Journal of Economic Theory 22, 477–498.

Hirshleifer, D., Lim, S. S., Teoh, S. H., 2009. Driven to distraction: Extraneous events and underreaction to earnings news. Journal of Finance 64, 2289–2325.

Hoberg, G., Lewis, C., 2017. Do fraudulent firms produce abnormal disclosure? Journal of Corporate Finance 43, 58–85.

Hughes, J. S., Liu, J., Liu, J., 2007. Information asymmetry, diversification, and cost of capital. The Accounting Review 82, 705–729.

Ingram, R. W., Raman, K. K., Wilson, E. R., 1989. The information in governmental annual reports: A contemporaneous price reaction approach. Accounting Review pp. 250–268.

Kelley, E. K., Tetlock, P. C., 2013. How wise are crowds? Insights from retail orders and stock returns. The Journal of Finance 68, 1229–1265.

Lambert, R., Leuz, C., Verrecchia, R. E., 2007. Accounting information, disclosure, and the cost of capital. Journal of Accounting Research 45, 385–420.

Lawrence, A., 2013. Individual investors and financial disclosure. Journal of Accounting and Economics 56, 130–147.

Longstaff, F. A., Mithal, S., Neis, E., 2005. Corporate yield spreads: Default risk or liquidity? new evidence from the credit default swap market. The Journal of Finance 60, 2213–2253.

Lopez-Lira, A., 2023. Risk factors that matter: Textual analysis of risk disclosures for the cross-section of returns. Working paper. Jacobs Levy Equity Management Center for Quantitative Financial Research.

Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-Ks. The Journal of Finance 66, 35–65.

Loughran, T., McDonald, B., 2014a. Measuring readability in financial disclosures. The Journal of Finance 69, 1643–1671.

Loughran, T., McDonald, B., 2014b. Regulation and financial disclosure: The impact of plain English. Journal of Regulatory Economics 45, 94–113.

Loughran, T., McDonald, B., 2020. Textual analysis in finance. Annual Review of Financial Economics 12, 357–375.

Miller, B. P., 2010. The effects of reporting complexity on small and large investor trading. The Accounting Review 85, 2107–2143.

Moss, A., Naughton, J. P., Wang, C., 2023. The irrelevance of environmental, social, and governance disclosure to retail investors. Management Science, forthcoming.

Painter, M., 2020. An inconvenient cost: The effects of climate change on municipal bonds. Journal of Financial Economics 135, 468–482.

Reck, J. L., Wilson, E. R., 2006. Information transparency and pricing in the municipal bond secondary market. Journal of Accounting and Public Policy 25, 1–31.

Schultz, P., 2013. State taxes, limits to arbitrage and differences in municipal bond yields across states. Working paper. University of Notre Dame.

Veldkamp, L. L., 2011. Information choice in macroeconomics and finance. Princeton University Press.

Verrecchia, R. E., 1982. Information acquisition in a noisy rational expectations economy. Econometrica 50, 1415–1430.

Wang, S., Liu, X., Zhou, J., 2022. Readability is decreasing in language and linguistics. Scientometrics 127, 4697–4729.

Wasike, B., 2016. Preaching to the choir? An analysis of newspaper readability vis-a-vis public literacy. Journalism 19, 1570–1587.

White, M., 2016. A new model for SEC enforcement: Producing bold and unrelenting results. https://www.sec.gov/news/speech/chair-white-speech-new-york-university-111816.

WSJ, 2019. Index funds are the new kings of Wall Street. https://www.wsj.com/articles/index-funds-are-the-new-kings-of-wall-street-11568799004.

WSJ, 2021. It isn't just AMC. Retail traders increase pull on the stock market. https://www.wsj.com/articles/it-isnt-just-amc-retail-traders-increase-pull-on-the-stock-market-11624008602.

WSJ, 2022a. Bed Bath & Beyond's stock-sale deluge lures retail investors. https://www.wsj.com/articles/bed-bath-beyonds-stock-sale-deluge-lures-retail-investors-bb9db1c5.

WSJ, 2022b. Local governments, many on Wall Street line up against muni-data bill. https://www.wsj.com/articles/local-governments-many-on-wall-street-line-up-against-muni-data-bill-11669393779.

You, H., Zhang, X. J., 2009. Financial reporting complexity and investor underreaction to 10-K information. Review of Accounting Studies 14, 559–586.

Figure 1: Excerpts from Official Statements. Panel A provides an excerpt from a municipal bond official statement with below-median complexity, and Panel B provides an excerpt from a municipal bond official statement with above-median complexity.

Panel A: Official Statement with Below-Median Complexity

TALLADEGA COUNTY, ALABAMA SPECIAL OBLIGATION SCHOOL WARRANTS, SERIES 2009 (AL) (Excerpt from Page 15)

Capital Project Fund

The Capital Project Fund is used to account for financial resources to be used for the acquisition or construction of major capital facilities (other than those financed by proprietary and trust fund). Fiduciary funds are used to account for assets held on behalf of outside parties including other governments, or on behalf of other funds within the government. When these assets are held under the terms of a formal trust agreement, either a nonexpendable trust fund or an expendable trust fund is used. The terms "nonexpendable" and "expendable" refer to whether or not the government is under an obligation to maintain trust principal. Agency funds generally are used to account for assets that the government holds on behalf of other as their agent. The Board's Agency Fund includes assets and related liabilities under a deferred compensation plan.

Account groups are used to establish accounting control and accountability for the Board's general fixed assets and the unmatured principal of its general long-term debt. These account groups are not funds. They do not reflect available financial resources and related liabilities but are accounting records of the general fixed assets and general long-term debt and certain associated information. Basis of Accounting. The basis of accounting refers to when revenues and expenditures or expenses are recognized in the account and reported in the financial statements.

Complexity Metrics
Page count: 144
Word count: 52,061
Readability grade level: 11.8
Complexity z-score: -0.38

Panel B: Official Statement with Above-Median Complexity

MASSACHUSETTS HEALTH AND EDUCATIONAL FACILITIES AUTHORITY REVENUE BONDS, MASSACHUSETTS EYE AND EAR INFIRMARY ISSUE SERIES C (2010) (MA) (Excerpt from Page C-5-3)

Application of Moneys

If available moneys in the Debt Service Fund after any required transfers from the Debt Service Reserve Fund and Redemption Fund are not sufficient on any day to pay all principal (including sinking fund installments), redemption price and interest on the Outstanding Bonds then due or overdue, such moneys (other than any sum in the Redemption Fund irrevocably set aside for the redemption of particular Bonds or required to purchase Bonds under outstanding purchase contracts) shall, after payment of all charges and disbursements of the Trustee in accordance with the Agreement, be applied (in the order such Funds are named in this section) first to the payment of interest, including interest on overdue principal, in the order in which the same became due (pro rata with respect to interest which became due at the same time), and second to the payment of principal (including sinking fund installments) and redemption premiums, if any, without regard to the order in which the same became due (in proportion to the amounts due). For this purpose interest on overdue principal shall be treated as coming due on the first day of each month. Whenever moneys are to be applied pursuant to this paragraph, such moneys shall be applied at such times, and from time to time, as the Trustee in its discretion shall determine, having due regard to the amount of such moneys available for application and the likelihood of additional moneys becoming available for such application in the future. Whenever the Trustee shall exercise such discretion it shall fix the date (which shall be the first of a month unless the Trustee shall deem another date more suitable) upon which such application is to be made, and upon such date interest on the amounts of principal paid on such date shall cease to accrue. The Trustee shall give such notice as it may deem appropriate of the fixing of any such date. When interest or a portion of the principal is to be paid on an overdue Bond, the Trustee may require presentation of the Bond for endorsement of the payment.

Complexity Metrics
Page count: 190
Word count: 96,825
Readability grade level: 23.6
Complexity z-score: 1.2

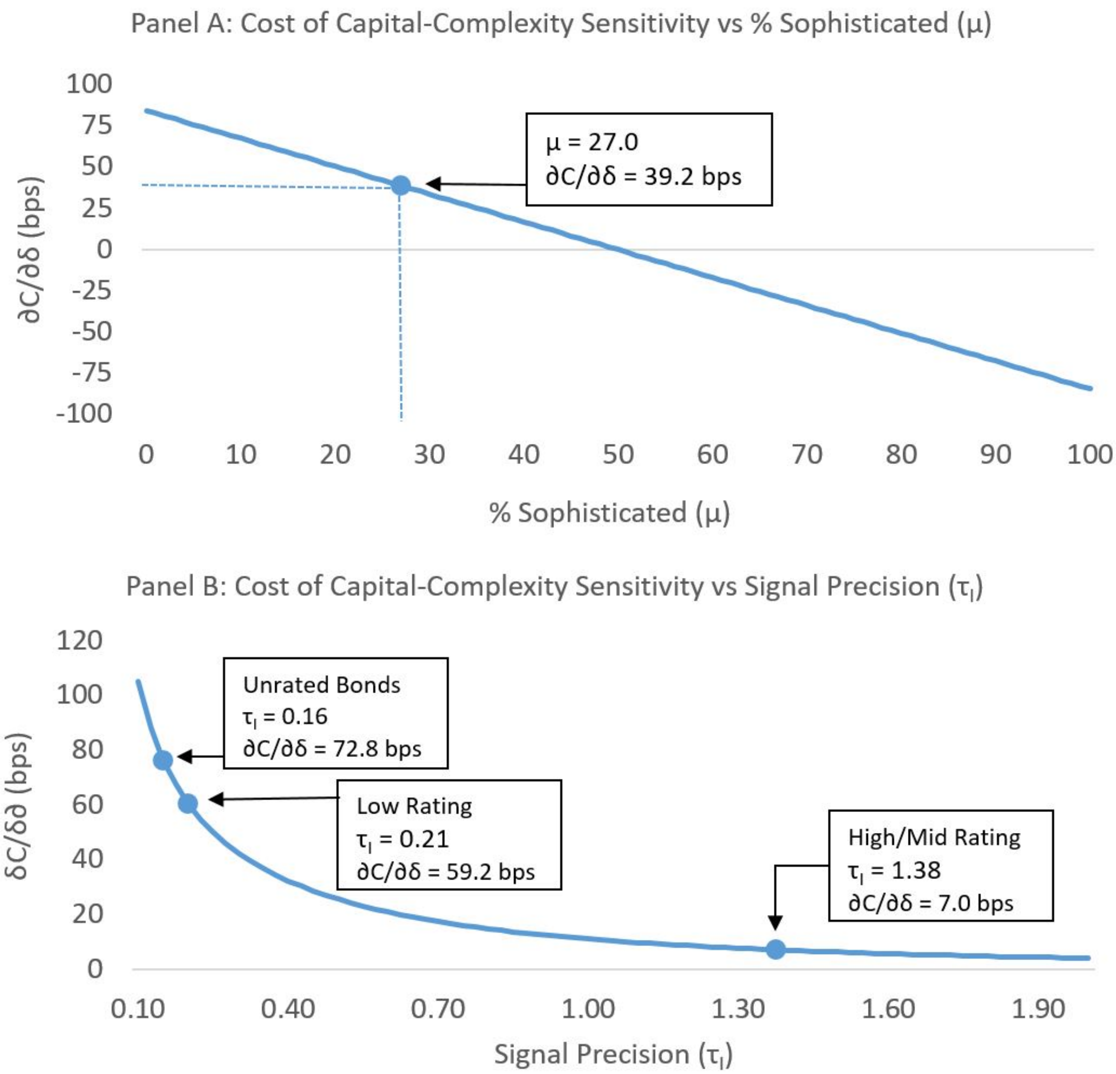
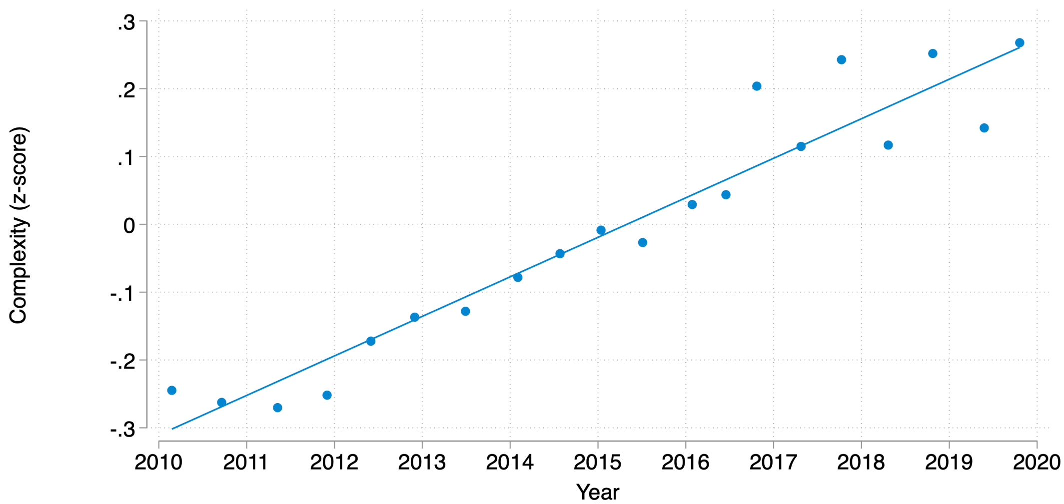


Figure 2: Cost of Capital-Complexity Sensitivity. Panel A graphs the cost of capital-complexity sensitivity versus the proportion of sophisticated traders (μ). Cost of capital-complexity sensitivity is calculated as $(\partial C/\partial \delta) \times (0.35/100)$, where the expression for $\partial C/\partial \delta$ is found in equation (9), 0.35 represents a one standard deviation change in δ, and 100 is a normalization term for a municipal bond with face value \$100. Cost of capital-complexity sensitivity is expressed in basis points. Panel A uses the following input parameters: $Q = 1$, $\delta = 0.5$, $\gamma = 1$, $\tau_x = 10$, and $\tau_I = \tau_J = \tau_v = 0.33$. Panel B graphs cost of capital-complexity sensitivity versus signal precision (τ_I) and uses the following input parameters: $Q = 1$, $\delta = 0.5$, $\gamma = 1$, $\tau_I = \tau_J = \tau_v$, $\tau_x = 10$, and $\mu = 0.27$.

Panel A: Raw Tone and Complexity Measures Over Time



Panel B: Residualized Tone and Complexity Measures Over Time

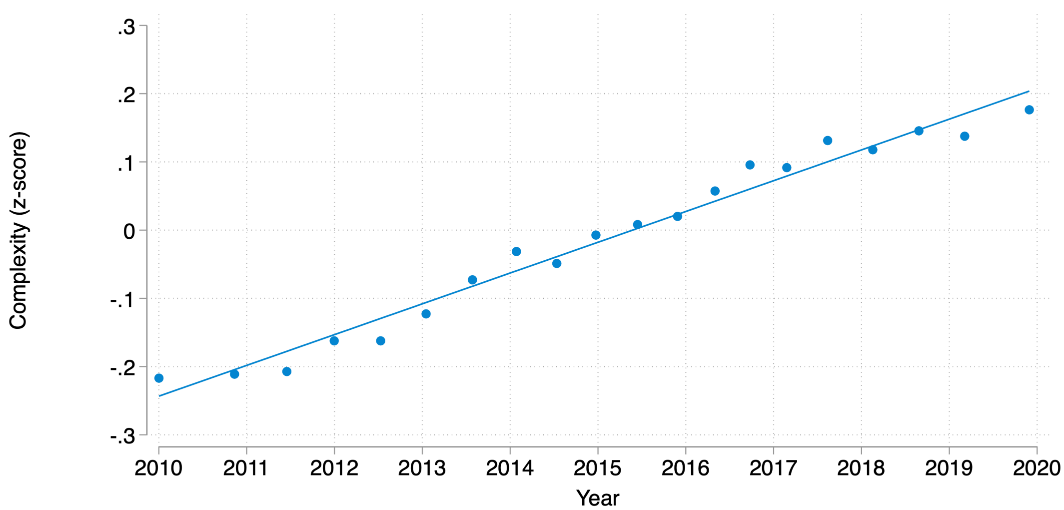


Figure 3: Time Series of Complexity in Official Statements. Panel A reports average complexity for each year during our sample period of mid-2009 to mid-2020. Panel B reports average residual complexity for each year during our sample period, where residual complexity is obtained from a regression of complexity on the control variables used in our baseline specification in equation (12).

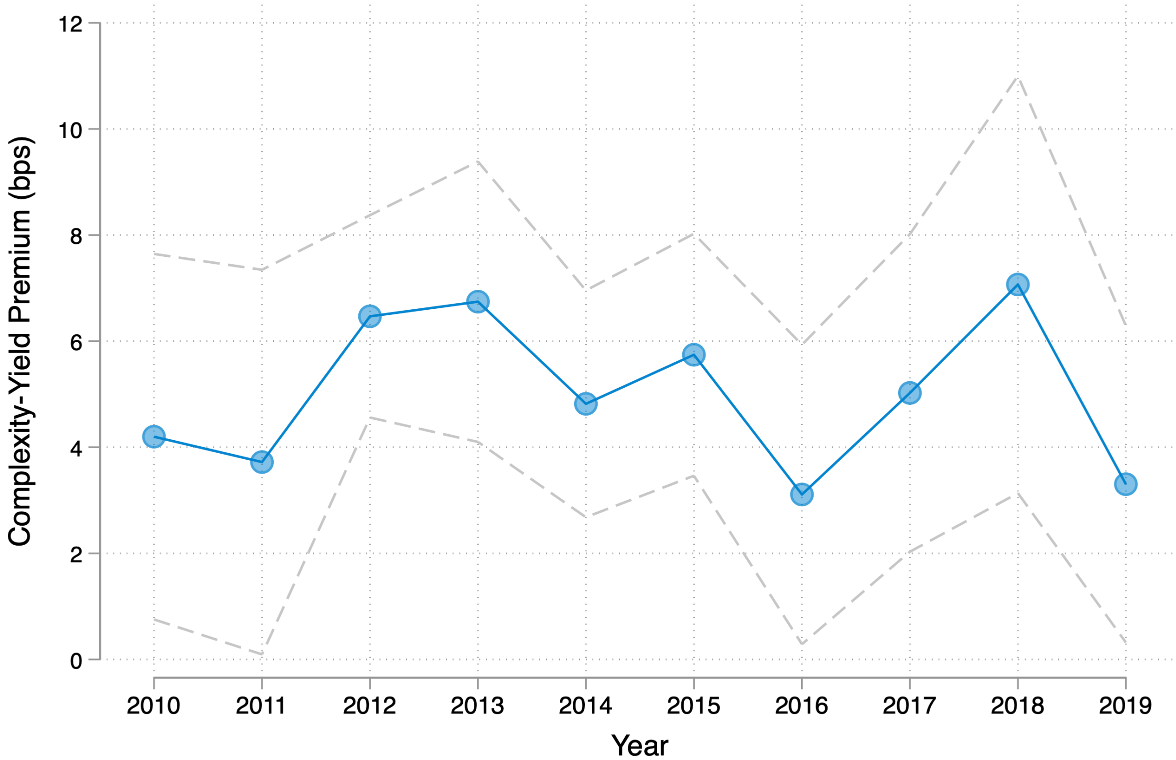


Figure 4: Complexity-Yield Premium by Year. This graph plots the complexity-yield premium for each full year of our sample period. The complexity-yield premium for each year is obtained by estimating our baseline regression model with the inclusion of interactions of complexity and an indicator variable for each year. The grey dashed lines represent the upper and lower bounds for the 95% confidence interval of the complexity-yield premium each year.

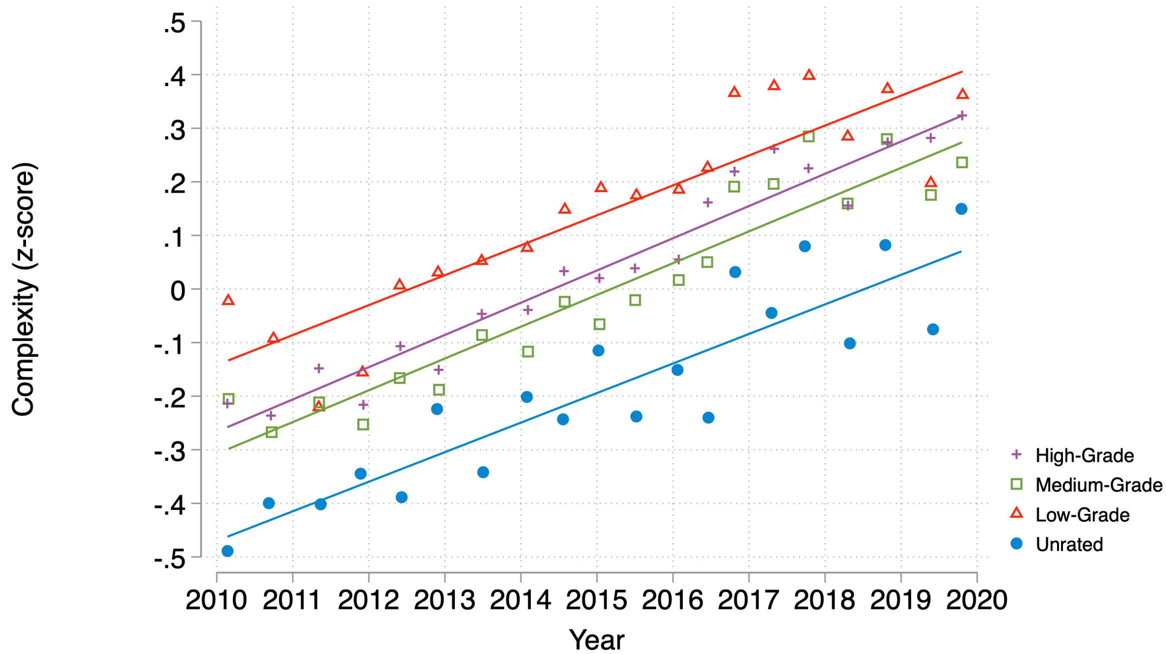


Figure 5: Complexity by Rating Category. This figure presents linear trends in municipal bond official statement complexity during our sample period for the following rating categories: high-rated bonds, medium-rated bonds, low-rated bonds, and unrated bonds.

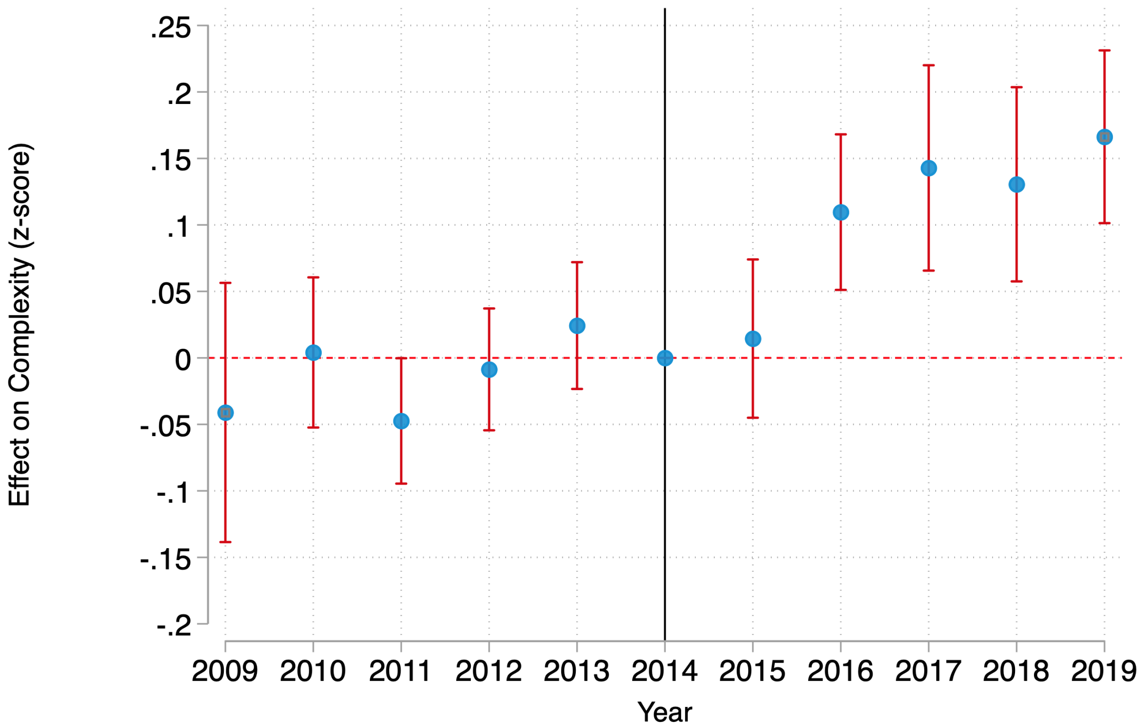


Figure 6: Effect of Regulatory Burden on OS Complexity. This figure shows the evolution of complexity for states that experience above-median growth in SEC enforcement officers (HEG states) relative to states that experience below-median growth during the SEC enforcement push from 2015 to 2017.

Table 1: Complexity and Bond Issue Summary Statistics. Panel A provides complexity-related summary statistics for our sample of municipal bond official statements. Number of words and number of characters are expressed in thousands (K). Flesch-Kincaid Grade and Gunning Fog Grade represent the estimated education grade level that is needed to understand the official statement. Panel B presents bond issue summary statistics across our sample of municipal bond issues. For each issue observation, we first calculate the size-weighted average offering yield spread, issue size, and years to maturity across the bonds within that issue. All remaining variables are defined in the main text.

Panel A: Complexity Summary Statistics

	Mean	Median	P25	P75	SD
Number of Pages	135.4	118.0	81.0	169.0	89.8
Number of Words (K)	63.4	48.8	26.8	80.3	57.3
Number of Characters (K)	260.4	214.6	124.0	329.9	200.0
Flesch-Kincaid Grade	20.2	18.9	16.1	22.7	6.6
Gunning Fog Grade	21.0	19.7	17.0	23.5	6.7
Observations	85,317				

Panel B: Bond Issue Summary Statistics

	Mean	Median	P25	P75	SD
Offering Yield Spread (%)	0.354	0.247	-0.010	0.599	0.623
Issue Size (M)	32.289	8.000	3.275	21.545	125.044
Years to Maturity	8.281	8.055	5.036	10.686	5.267
Rating Number	18.359	19.000	17.000	20.000	1.971
Unrated	0.257	0.000	0.000	1.000	0.437
General Obligation	0.612	1.000	0.000	1.000	0.487
Insured	0.128	0.000	0.000	0.000	0.334
Callable	0.392	0.000	0.000	1.000	0.488
Negotiated	0.332	0.000	0.000	1.000	0.471
Bank Qualified	0.499	0.000	0.000	1.000	0.500
Observations	85,317				

Table 2: Complexity and Yield Spread. Each numbered column in this table reports the results of an OLS regression of municipal bond yield spread on complexity. Column (2) includes municipal bond and county-level control variables, and columns (3) and (4) also include state fixed effects and state-year fixed effects respectively. Standard errors are double clustered by county and year-month, and are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Dep. Variable: Yield Spread			
	(1)	(2)	(3)	(4)
Complexity	0.101***	0.034***	0.042***	0.049***
	(0.013)	(0.009)	(0.009)	(0.008)
Issue Size		-0.059***	-0.052***	-0.054***
		(0.007)	(0.007)	(0.007)
Maturity		0.187***	0.191***	0.192***
		(0.021)	(0.022)	(0.023)
Negotiated		0.110***	0.109***	0.117***
		(0.016)	(0.016)	(0.016)
Gen. Obligation		-0.298***	-0.297***	-0.295***
		(0.015)	(0.016)	(0.015)
Insured		0.037**	0.009	0.033**
		(0.016)	(0.015)	(0.015)
Callable		0.061***	0.063***	0.050***
		(0.017)	(0.017)	(0.016)
ln(Population)		0.038***	0.034***	0.027***
		(0.007)	(0.006)	(0.006)
Unemp. Rate		0.054***	0.061***	0.020***
		(0.007)	(0.008)	(0.003)
Fixed Effects	None	None	State	State-Year
R^2	0.022	0.243	0.272	0.345
N	84,546	84,545	84,545	84,542

Table 3: Complexity and Yield Spread across Rating Categories. Each numbered column in this table reports the results of an OLS regression of municipal bond yield spread on complexity within a particular rating categories. The "High" and "Medium" rating subsamples in columns (1) and (2) represent the municipal bonds that are placed in the highest two credit rating categories and next-highest two credit rating categories, respectively. The "Low" rating subsample in column (3) represents the municipal bonds that are placed in all other credit rating categories. The "Unrated" category in column (4) represents the municipal bonds that are not rated by any credit rating agency. Standard errors are double clustered by county and year-month, and are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Yield Spread	(2) Yield Spread	(3) Yield Spread	(4) Yield Spread
Complexity	0.013***	0.006*	0.074***	0.091***
	(0.004)	(0.003)	(0.008)	(0.017)
Rating	High	Medium	Low	Unrated
Controls	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.364	0.472	0.486	0.631
N	18,071	28,188	16,693	21,484

Table 4: Complexity and Yield Spread for General Obligation and Revenue Bonds. Each numbered column in this table reports the results of an OLS regression of municipal bond yield spread on complexity among GO and revenue bonds. The regression in column (1) focuses on the subsample of GO bonds. The regressions in columns (2), (3), and (4) focus on the subsample of revenue bonds that have a high/medium credit rating, a low credit rating, and no credit rating, respectively. Standard errors are double clustered by county and year-month, and are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Yield Spread	(2) Yield Spread	(3) Yield Spread	(4) Yield Spread
Complexity	-0.024***	0.037***	0.105***	0.140***
	(0.005)	(0.005)	(0.011)	(0.019)
GO	Yes	No	No	No
Rating	All	High/Medium	Low	Unrated
Controls	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.391	0.422	0.466	0.621
N	52,036	15,974	8,810	7,615

Table 5: Complexity and Yield Spread by Tax Privilege. Column (1) reports the results of an OLS regression of municipal bond yield spread on complexity for the subsample of states without tax privileges for municipal bond income; column (2) analogously reports for the states with such tax privileges. In column (3), we use the full sample in regression of yield spread on complexity: the interaction of complexity with an indicator variable that equals one if the bond was issued in a state with no tax privilege. In column (4), we use the full sample and regress yield spread on complexity: the interaction of complexity with an indicator variable that equals one if the bond was issued in a state that had no tax privilege in 1947. Standard errors, double clustered by county and year-month, are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Yield Spread	(2) Yield Spread	(3) Yield Spread	(4) Yield Spread
Complexity	0.023	0.056***	0.061***	0.067***
	(0.015)	(0.010)	(0.010)	(0.011)
Comp. × No Tax Priv.			-0.052**	
			(0.024)	
Comp. × No Priv. (1947)				-0.052**
				(0.021)
State	No Tax Priv.	Tax Priv.	All	All
Controls	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.306	0.365	0.346	0.346
N	22,006	62,536	84,542	84,542

Table 6: Complexity and Yield Spread by Bond Characteristic. Each numbered column in this table reports the results of an OLS regression of municipal bond yield spread on complexity with interactions between complexity and bond characteristics. In column (1), complexity is interacted with a "High Price" indicator which equals one if the bond issuance price is at least $103. In column (2), complexity is interacted with a "BQ" indicator which equals one if the bond is bank qualified. In column (3), complexity is interacted with a "Negotiated" indicator, which equals one if the bond was issued through a negotiated sale. In column (4), we include all of the interactions from the first three numbered columns. Standard errors are double clustered by county and year-month, and are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Yield Spread	(2) Yield Spread	(3) Yield Spread	(4) Yield Spread
Complexity	0.066***	0.091***	0.071***	0.105***
	(0.011)	(0.013)	(0.014)	(0.017)
Complexity × BQ	-0.067***			-0.084***
	(0.013)			(0.014)
Complexity × High Price		-0.096***		-0.113***
		(0.013)		(0.014)
Complexity × Negotiated			0.070***	0.043***
			(0.017)	(0.012)
Controls	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.357	0.395	0.238	0.409
N	84,542	70,902	84,543	70,902

Table 7: Complexity and Yield Volatility. Column (1) reports the results of an OLS regression of municipal bond yield volatility on complexity. The remaining numbered columns include one of more complexity interaction terms with complexity. In column (2), we interact complexity with the medium credit rating, low credit rating, and unrated indicator variables. In columns (3) and (4), we interact complexity with indicators for GO and high price, respectively. Yield volatility is calculated as the size-weighted average standard deviation in secondary yield spread across bonds within each issue. Standard deviation is calculated using the three-year window after issuance. Standard errors are double clustered by county and year-month, and are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Yield Vol.	(2) Yield Vol.	(3) Yield Vol.	(4) Yield Vol.
Complexity	0.003***	-0.002	0.006***	0.008***
	(0.001)	(0.001)	(0.002)	(0.003)
Comp. × Mid Rating		0.000		
		(0.001)		
Comp. × Low Rating		0.010***		
		(0.003)		
Comp. × Unrated		0.022***		
		(0.006)		
Comp. × GO			-0.006**	
			(0.003)	
Comp. × High Price				-0.008***
				(0.003)
Controls	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.113	0.125	0.114	0.112
N	53,752	53,752	53,752	45,965

Table 8: Complexity and Liquidity. Column (1) reports the results of an OLS regression of issue-level average roundtrip transaction cost (Markup) on complexity. Columns (2) and (3) repeat the same regression, except with issue-level average retail roundtrip transaction cost (RMarkup) and issue-level average institutional roundtrip transaction cost (RMarkup), respectively. Retail transactions are defined as transactions that are less than $100 thousand in trade size; institutional transactions are defined as transactions that are greater or equal to $100 thousand in trade size. Roundtrip transaction costs are calculated using a three-year window after each issuance. Standard errors are double clustered by county and year-month, and are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Markup	(2) Ret. Markup	(3) Inst. Markup
Complexity	0.012***	0.017***	-0.011**
	(0.004)	(0.005)	(0.005)
Controls	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes
R^2	0.454	0.281	0.194
N	27,867	27,867	27,867

Table 9: LDA Topics Mapping. This table presents the top three representative words for select LDA-generated topics. We assign these words to one of the following four categories: Risk, Bond Description, Cash Flow, and Legal.

Risk
1. Security, Risk, Fund
2. Liability, Deferred, Resource
3. Default, Event, Trustee

Bond Description
1. Redemption, Bond, Date
2. Project, Improvement, Capital
3. Bond, Price, Discount

Cashflow
1. Budget, Appropriation, Year
2. Expense, Fee, Cost
3. Fund, Investment, Portfolio

Legal
1. Mean, Resolution, Shall
2. Bond, Opinion, Counsel
3. Authority, Finance, Resolution

Table 10: Topic Complexity and Yield Spread by Rating Category. Each column reports the results of an OLS regression of municipal bond yield spread onto topic complexities. We first use the LDA model to subdivide the OS into components based on four broad categories: Risk, Legal, Cash Flow, and Bond Description. We then calculate the document's complexity based on the four components. In columns (1) to (4), we focus on the subsamples of high-rated bonds, medium-rated bonds, low-rated bonds, and unrated bonds, respectively. Standard errors are double clustered by county and year-month, and are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Yield Spread	(2) Yield Spread	(3) Yield Spread	(4) Yield Spread
Risk Complexity	0.003	-0.000	0.052***	0.043***
	(0.006)	(0.005)	(0.010)	(0.014)
Legal Complexity	0.045***	0.033***	0.123***	0.198***
	(0.006)	(0.006)	(0.012)	(0.020)
Cash Flow Complexity	0.005	-0.001	-0.008	-0.004
	(0.005)	(0.005)	(0.008)	(0.009)
Description Complexity	0.004	0.013***	0.023***	0.055***
	(0.005)	(0.004)	(0.008)	(0.012)
Rating	High	Medium	Low	Unrated
Controls	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.369	0.474	0.504	0.655
N	18,071	28,188	16,693	21,484