

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

Working Paper

The complexity yield puzzle: A textual analysis of municipal bond disclosures

New Working Paper Series, No. 338

Provided in Cooperation with:

George J. Stigler Center for the Study of the Economy and the State, The University of Chicago Booth School of Business

Suggested Citation: Farrell, Michael; Murphy, Dermot; Painter, Marcus; Zhang, Guangli (2023) : The complexity yield puzzle: A textual analysis of municipal bond disclosures, New Working Paper Series, No. 338, University of Chicago Booth School of Business, Stigler Center for the Study of the Economy and the State, Chicago, IL

This Version is available at:

<https://hdl.handle.net/10419/279822>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Stigler Center

for the Study of the Economy and the State

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

Michael Farrell

University of Wisconsin-Milwaukee

Dermot Murphy

University of Illinois at Chicago

Marcus Painter

Saint Louis University

Guangli Zhang

Saint Louis University

November 2023

New Working Paper Series No. #338

Stigler Center for the Study of the Economy and the State

University of Chicago Booth School of Business

5807 S Woodlawn Ave

Chicago, IL 60637

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

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

November 3, 2023

Abstract

We study the effect of official statement textual complexity on municipal borrowing costs (yield) through the clientele channel. Theoretically, complexity reduces signal quality for unsophisticated investors, potentially increasing yield. We provide causal evidence of an economically significant complexity-yield premium that is especially large for issues with greater tax-induced exogenous demand from unsophisticated investors or less complementary information from rating agencies. 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. This trend is attributable to increased regulatory enforcement from the SEC.

JEL classification: G12, G14, G18, H74

Keywords: municipal bonds, disclosure, complexity, textual analysis, clientele

*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 Bergstresser, Chris Berry, Igor Cunha, Christine Cuny, David Erdman, Dan Garrett, Ralph Koijen, Tim Loughran, Lori Raineri, Win Smith, and seminar participants at CUNY-Queens, Oklahoma State University, Saint Louis University, Southern Illinois University, the University of Illinois Chicago, the University of Nevada, Reno, the Virtual Municipal Finance Workshop, the 12th Annual Brookings Municipal Finance Conference, 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 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 US municipal bond market, where direct participation by individual investors from modest income brackets is high relative to other markets (Bergstresser, 2023; *Municipal Bonds for America*, 2016), 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 and a lack of common structure for the accompanying official statements, unlike equity or corporate bond markets. 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. Furthermore, 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 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 decreases signal precision for unsophisticated investors and increases it for sophisticated 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 into our theoretical model to infer the overall clientele makeup in this 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 makeup 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 close to our model-implied estimate.

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.² Our first proxy is the tax privilege provided to investors for

¹See [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 \(1982\)](#).

²Retail investors are not necessarily unsophisticated and thus incapable of processing complex informa-

holding local municipal bonds, which is known to increase the proportion of retail investors that are active in the local market (Babina et al., 2021). Our second and third proxies are the bank qualification status and the offering price premium on the issue, 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 issues that are more likely to have a larger retail investor base. We also find the complexity-yield premium is significantly larger for bonds issued in the negotiated market, where retail investors are often given a priority purchase window at offering. 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 cases where official statement complexity 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 issuance quality. Similarly, official statement complexity is less likely to matter for general obligation (GO) issues because potential investors already have access to historical economic information from county government websites and other sources, unlike revenue bond issues which frequently rely on a specific revenue source with limited history. The evidence we find is consistent with these conjectures. In particular, the complexity-yield premium for unrated bonds (9.1 basis points) is almost twice as large as the average premium (4.9 basis points). Furthermore, we find no evidence of a positive complexity-yield premium for GO bonds, and strong evidence of a 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.

One potential concern with our central findings is that complexity and offering yield spread are jointly influenced by a latent variable that is not captured by our battery of control variables (which includes proxies for the instrumental complexity of the issue) and fixed effects. To take a further step toward a causal interpretation, we employ a spatial “leave one out” instrumental variable (IV) approach that exploits the within-state average

tion. 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.

complexity of other negotiated issue official statements the underwriter has helped write. Conversations with industry professionals support the notion that the writing style and legal language used by an underwriter are plausibly exogenous to the traits of a bond issuance. The IV results continue to show a complexity-yield premium that is concentrated in low and unrated issues, suggesting a causal interpretation of our results.

Given that official statement complexity is costly for issuers, one might expect to see a simplification of official statements over time. However, our time-series analysis shows that official statement complexity has increased steadily over our sample period. In particular, we find that the average numbers of pages and 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 1.8. Broadly, these changes represent a 0.50 standard deviation increase in our complexity measure. This upward trend in complexity does not appear to be driven by the 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 the remainder of our analysis explores reasons for this puzzle.

We consider two non-mutually exclusive hypotheses that may explain the complexity-yield puzzle. The catering hypothesis posits that issuers write complex statements in an effort to attract the growing 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 add more information to official statements in response to heightened regulatory oversight (Dyer et al., 2017). Our evidence primarily supports the regulatory burden hypothesis, in that official statements have become significantly more complex in states that experienced a relatively higher surge in local Securities and Exchange Commission (SEC) enforcement officers during a major enforcement push in the mid-2010s. There is less support for the catering hypothesis, as official statement complexity has uniformly increased across bond types, including those with more institutional investors such as bank-qualified bonds and high-premium bonds. Our estimate of the complexity-yield premium is also stable over time, suggesting that issuers have not effectively targeted institutions to reduce the complexity-yield premium. However, we do find that official statement complexity has trended more strongly in states that have no local tax privilege and thus a larger institutional clientele makeup, lending some support to the catering hypothesis.

To further investigate the factors influencing complexity, we employ a machine learning technique that enables us to identify 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 issuers 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 categories from the LDA topics – risk, legal, cash flow, and description – and measure complexity within each category. We find that the complexity-yield premium is primarily driven by legal complexity, which further supports our regulatory burden hypothesis.

Our study 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 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 information processing costs are high ([Hirshleifer et al., 2009](#)). Our study bridges the gap between the literature 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.

Second, our study 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](#)). (See [Loughran and McDonald, 2020](#) for a recent review of other work in this area.) 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). Ertugrul et al. (2017) show that length and ambiguity in annual reports are associated with higher corporate 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 illustrating a strong causal effect of complex disclosure on the cost of capital that hinges on the composition of investor clientele.

Third, our study contributes to the collective knowledge on relevant determinants of pricing and trading activity in the municipal bond market. Empirical evidence is mixed on whether information disclosures from issuers affect trading activity in the secondary market. Ingram et al. (1989) and Reck and Wilson (2006) find no evidence that government annual reports influence municipal bond prices. However, 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).³ Our findings shed light on this question by showing that official statement complexity does influence the pricing of municipal bonds, especially those with a large retail investor presence, which in turn affects local economic conditions (Adelino et al., 2017).⁴ Our in-depth textual analysis of official statements also allows us to add nuance to the debate on the relevance of municipal disclosures that could not otherwise be achieved by looking solely at the release of disclosure documents. Lastly, 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

³In the corporate setting, Loughran and McDonald (2017) find that the number of daily requests for corporate 10-K filings through the SEC's EDGAR system is surprisingly low, suggesting that many investors are not doing fundamental research on stocks. Given this evidence, many municipal bond investors may not be accessing official statements either. Municipal bond investors also rely greatly on credit ratings to assess bond quality (Cornaggia et al., 2018), which further raises the question of whether primary market disclosures are referenced by many investors. Our later findings illustrate that the complexity-yield premium is especially large for unrated and low-rated bonds, where credit rating reports are less likely to act as a good substitute for the official statement.

⁴Other recent work has found that municipal bond yields and trading outcomes are affected by the timeliness of trade reports (Chalmers et al., 2021), pension underfunding (Novy-Marx and Rauh, 2012; Giesecke et al., 2023), 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).

[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.⁵ 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](#)).

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 fragmented with over one million outstanding bonds issued by approximately 50,000 state and local governments. Bond issuances finance a variety of projects that are supported by the local tax base (GO bonds) or the cash flows generated by the underlying project (revenue bonds). On the demand side, retail investors comprise the largest category of municipal bond investors, holding just under half of outstanding bonds by par value. This is especially large among fixed income asset classes; for comparison, households hold only 3% of US Treasuries and 1% of corporate bonds ([Bagley et al., 2023c](#)). Part of the reason for the high retail participation is that municipal bond interest income is typically exempt from state taxes if the bond was issued in that state, thereby motivating local investors to purchase in-state municipal bonds.

Retail investors are active on the primary and secondary markets. On the primary market, they 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 risks 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

⁵Related 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, 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).

Retail investors in the municipal bond market are also more likely to come from modest income tax brackets and less likely to use a financial advisor, underlining the notion that retail investors are unsophisticated relative to institutional investors and have fewer resources to process complex official statements. Northwestern Mutual (2022) and NerdWallet (2020) note that the majority of Americans do not use a financial advisor. Although advisor use may be higher among investors with higher incomes, we find that a large share of municipal bond holders do not have incomes at the top end of the income distribution. Using data from the Survey of Consumer Finances (SCF) during the 2010 to 2019 period, we find that the median household income for a municipal bond holder is \$282,000, and that the 25th percentile of household income is only \$100,000. We also find that 40% of municipal bond holders do not use a financial planner in their investment decision making process. Lastly, there is little evidence that municipal bond ownership is dominated by high income households on the intensive margin, as 40% of municipal bond interest is paid to households with an annual income below \$200,000 (Municipal Bonds for America, 2016). Overall, this analysis suggests that the burden of processing municipal bond disclosures falls directly on a large share of retail investors with modest resources for processing complex statements.

Institutional investors are also highly active in the municipal bond market and hold the remaining 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 REE 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. 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 ([Schultz, 2013](#); [Babina et al., 2021](#)) that may not have the resources to process large, complex statements.

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 $\tau_v > 0$, where precision is the reciprocal of variance. In other words, we have that $v \sim N(0, \tau_v^{-1})$. The 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 $i \in I$ receives a signal s_i about the terminal payoff v of the risky municipal bond that equals:

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

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

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

Importantly, the δ parameter measures the level of disclosure that is required for the official statement. 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 D 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:

$$p = p_0 + p_v v + p_x x, \quad (3)$$

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 has mean-variance utility over

their final wealth, and thus their demand D for the municipal bond takes the following form:

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

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:

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

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

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

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:

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

Then, 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 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:

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

Lastly, we note that $B > 0$, which implies that F , 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 in our model, we have a unique and partially revealing noisy REE.

Importantly, we can now determine how an exogenous increase in disclosure complexity δ affects three equilibrium market 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 ultimately absorb liquidity demand through the dealer market. Lastly, our third outcome is the 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:

$$\partial C/\partial\delta = -\frac{C^2 \times (2\tau_x B + \gamma^2)}{Q\gamma} \times (\tau_I\mu - \tau_J(1 - \mu)) \quad (9)$$

$$\partial L/\partial\delta = \frac{L}{\gamma F} \times \frac{(\tau_x B + \gamma^2)^2 - \tau_x \tau_v \gamma^2}{\tau_x B + \gamma^2} \times (\tau_I\mu - \tau_J(1 - \mu)) \quad (10)$$

$$\partial\sigma/\partial\delta = -\frac{\gamma^2(F\tau_x^2 B + (2\tau_x B + \gamma^2)K)}{F^3\sigma\tau_x\tau_v} \times (\tau_I\mu - \tau_J(1 - \mu)), \text{ where} \quad (11)$$

$$K = \tau_v(\tau_x\tau_v\gamma^2 + (\tau_x B + \gamma^2)^2) - F\tau_x > 0.$$

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. For the C and σ 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 $\tau_I\mu - \tau_J(1 - \mu)$. For the L 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, 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 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, a means of marketing the investment, and all material information pertinent to the bonds such as the underlying 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 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 into a machine readable format using the PDFTOHTML software program. This process results in text data for 12,416,759 pages across 87,795 official statements. The EMMA dashboard metadata 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 textual complexity measures: the number of pages in the official statement, the 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.⁶ While these measures vary in their construction, they are highly correlated and tend to agree in the rankings of complexity between documents. (Related, if we exclude some of these readability measures from our complexity index or use a simpler measure of complexity such as the natural log of word count, then all of our empirical results carry through.) To construct the complexity index, we first calculate the z-score for each measure so that it has a mean of zero and a standard deviation of one. The complexity index, denoted *Complexity*, is constructed as the simple average of these z-scores. For ease

⁶The Flesch-Kincaid grade level index, for example, is calculated as $0.39 \times (\text{total words}/\text{total sentences}) + 11.8 \times (\text{total syllables}/\text{total words}) - 15.59$. We calculate the readability measures using Textstat (Aggarwal, 2023). See Table 2 of Wang et al. (2022) for the exact formulas and linguistic features associated with each of these measures.

of interpretation in our later regression tests, this index is standardized to have a mean of zero and a standard deviation of one.

Figure 1, Panels A and B provide text samples from a below-median complexity statement from Alabama and an above-median complexity statement from Massachusetts, respectively. 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 that from Alabama.⁷

Table 1, Panel A presents summary statistics for our complexity index inputs. The average document is fairly advanced, reading at an approximately 21st grade level. 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 contains 135 pages and 63,000 words. This high volume of information, in combination with the high average grade level needed to process 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 the information 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 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

⁷Very high readability grade levels should be taken with a grain of salt. For example, the readability grade level of 23.6 in Figure 1, Panel B suggests that the average reader would need a four-year undergraduate degree, a two-year masters degree, and a six-year doctorate degree to understand the document. Our preferred interpretation of readability grades above 16 is that the average reader would need a four-year undergraduate degree to understand the document, and that any grade level above 16 represents an extra amount of time needed to fully process the document. In our empirical tests, interpretation of readability grade levels is unnecessary because our central complexity measure relies on z-score measures of readability grade levels in the cross-section of official statements.

is rated), callable status, funding source (GO 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 local economic variables can also influence our outcome variables. We also supplement the bond data with secondary market trading information from the MSRB EMMA database, which contains the universe of municipal bond trades. 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 MSRB EMMA platform, 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 to 60 basis points.⁸ A negative yield spread 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. GO bonds represent 61% of issues in the sample, and the remaining 39% of issues are revenue bonds. 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 often 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.

⁸The issue-level offering yield spread is calculated as follows. For each municipal bond, we first calculate its yield spread as the difference between its offering yield and the coupon-equivalent US Treasury yield. The coupon-equivalent US Treasury yield is calculated by taking the present value of the future bond payments using the zero-coupon US 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.

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 is the issue-level offering yield spread (y), and the central independent variable 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). According to our theoretical model, the effect of complexity on yield spread is unclear because it depends on whether or not the total 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:

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

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 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 in columns (2) to (4), we progressively include the control vector X , state-year fixed effects, and issuer fixed effects.⁹ Across all specifications, we find that there is a positive and statistically significant relationship between complexity and yield spread. Our preferred specification in column (3) indicates that a one standard deviation increase in complexity is associated with a 4.9 basis point increase in the yield spread, or a “complexity-yield premium” of 4.9 basis points. In relative terms, this 4.9

⁹An advantage of including issuer fixed effects in column (4) is that we effectively control for unobservable issuer-specific risks that may be driving complexity or obfuscation of true credit risk, although a disadvantage is that we lose information from small issuers that do not frequently appear in our sample and are more reliant on local retail investors for raising capital.

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). For a typical issue with an issuance size of \$32 million and a duration of eight years, the 4.9 basis point complexity-yield premium represents an equivalent “complexity-price discount” of 39.2 basis points, or \$125,000 in additional present value interest payments if the issue size is held constant. The 39.2 basis point discount 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 results indicate that complex official statements are associated with higher municipal borrowing costs that are economically significant for local municipalities.¹⁰

For illustrative purposes, we use the 39.2 basis point complexity-price discount 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 from the Securities Industry and Financial Markets Association (SIFMA) as a reality check to the model. 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 or 32% of bond market supply, where $(1/10)^{0.5} = 0.32$. $\tau_v = \tau_I = \tau_J = 0.33$ is equivalent to one standard deviation in municipal bond price (\$1.73), where $(1/0.33)^{0.5} = \$1.73$. $\delta = 0.5$ implicitly assumes that the sophisticated traders and unsophisticated traders have the same signal precision before complexity is taken into account.

Figure 2, Panel A provides a graph of our model-implied complexity-price discount ($\partial C/\partial \delta \times 0.35$) versus the proportion of sophisticated traders (μ).¹¹ First, we find that the complexity-price discount is positive as long as $\mu < 0.5$, and that the positive complexity-price discount gets larger as μ decreases. Intuitively, the complexity-price discount 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

¹⁰The control variable coefficients are largely expected and reflect findings in the prior literature. For example, the positive coefficient on the *Insured* indicator variable in the cross-issuer tests 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).

¹¹For normalization purposes, we multiply $\partial C/\partial \delta$ 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.

increase in complexity would increase the discount by 76 basis points. At the other extreme, if $\mu = 95\%$, then a one standard deviation increase in complexity would reduce the discount for municipal borrowers by 76 basis points.

Importantly, when $\mu = 27\%$, the corresponding complexity-price discount structural estimate of 39.2 basis points in Figure 2, Panel A is the same as the 39.2 basis point empirical estimate derived from our main regression specification in column (3) of Table 2. In reality, is this estimate of μ consistent with 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 close to our model implied estimate of 27%.

5.1 Complexity Effects in the Cross-Section of Risk

We next explore how the complexity-yield premium varies in the cross-section of municipal bond coverage and risk. Official statement complexity is less likely to matter if 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 act as 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 the reports may only act as partial substitutes for official statements. Therefore, we hypothesize that the complexity-yield premium is stronger for unrated and lower-rated municipal issues.

Formally, we re-test the regression model in equation (12) using the following 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, Panel A. In columns (1) and (2), we find that the complexity-yield premia for the high-quality and medium-quality groups are positive but small at 1.3 basis points and 0.6 basis points, respectively. By contrast, in column (3), we find a much larger complexity-yield premium 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 and difficult-to-understand cash flows. Lastly, in column (4), we find a similarly large complexity-yield premium of 9.1 basis points for unrated bonds, indicating that the premium is especially large when investors do not have access to credit rating reports.

In Figure 2, Panel B, we graph the structural complexity-price discount versus signal precision (τ_I) 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 discount 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 discount is decreasing in τ_I , indicating that the discount is larger for riskier bonds with underlying cash flows that are difficult to understand. Second, we can use our point estimates from Table 3, Panel A to back out structural estimates of τ_I . The complexity-yield premia documented in this table are equivalent to empirical complexity-price discounts 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.¹² These results underline the importance of high-quality information disclosure for reducing complexity-price discounts in the low-quality and unrated bond categories.

¹²This graph also allows us to explore variation in the complexity-price discount under different signal precision counterfactuals. For example, if low-rated or unrated issuers were able to increase the signal precision in their official statements to 0.40 (still well-below the average signal precision of 1.38 for higher-rated issuers), then the complexity-price discount significantly decreases to 32.2 basis points, or a 54.4% decrease for unrated bonds and a 44.2% decrease for low-rated bonds.

Lastly, we compare the complexity-yield premia for GO bonds and revenue bonds. GO 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, which is especially helpful when the official statement is complex or difficult to process. Information relevant to the GO bond market includes publicly-available historical financial records on county government websites, tax records, intergovernmental revenue transfers, and employment statistics. By contrast, revenue bonds are backed by cash flows from new projects which lack historical records or other information that would be helpful for interpreting complex official statements. We hypothesize that the complexity-yield premium will be more pronounced for revenue bonds, especially those that are low-rated or unrated and thus also 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 3, Panel B. In column (1), we find that the complexity-yield premium 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 premium 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 information outside what is supplied by credit rating agencies. Lastly, in columns (3) and (4), we find large complexity-yield premia 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

Our previous results indicate that there is an economically significant complexity-yield premium in the municipal bond market, especially if the associated issue lacks supporting information from credit rating agencies. Our model and structural estimates in Figure 2,

Panel A further indicate that the complexity-yield premium is theoretically increasing in the proportion of unsophisticated investors ($1 - \mu$), with the intuition that complexity reduces signal quality for unsophisticated investors, which in turn affects pricing outcomes. The purpose of this section is to empirically test if the complexity-yield premium varies with the proportion of unsophisticated investors in the market, thereby providing support for our proposed driver of the complexity-yield premium.

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 local municipal bonds. 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 above and also Alaska, Indiana, and Nevada, have significantly lower in-state ownership compared to the remaining four quintiles. Given that bonds issued in these “no tax privilege” states are less likely to be held by unsophisticated local investors, we hypothesize that the complexity-yield premium is lower in these 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 4. 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 premium in tax-privilege states is positive and statistically greater than the (statistically insignificant) premium in no-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 fewer 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 municipalities 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 \(2013\)](#), 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 5. 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 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 3.2 basis points for non-negotiated (competitive) bonds, and an even larger complexity-yield premium of 7.6 basis points for negotiated bonds. Lastly, in column (4), we include all of the variables from columns (1) to (3) and find similar results, although a small complexity-yield premium persists for BQ bonds (2.2 basis points). Overall, these results support our earlier tests on tax-induced clientele effects, and further highlight the importance of sophisticated investors for reducing the complexity-yield premium.

5.3 Underwriter Complexity Spatial Instrument

The negotiated bond subsample provides a unique opportunity to further establish a causal effect of official statement complexity on offering yield spread that does not rely on a cross-sectional comparison of bond types. For negotiated issues, the underwriter and the underwriter’s counsel are typically highly involved with drafting the official statement. According to municipal bond practitioners, the writing style of the official statement often depends on the writing style and legal language of the underwriter involved with the issue. Therefore, we use within-state average complexity of the associated underwriter as an instrument for

official statement complexity. For each issue observation, we calculate average complexity for all other official statements that are drafted by the same underwriter and issued in the same state, excluding any official statements from the same issuer. Formally, we employ a spatial “leave one out” IV regression which has been used in other studies such as [Acemoglu et al. \(2008\)](#) and [Persson and Tabellini \(2009\)](#), except that we exploit the identity of the underwriter to obtain an instrument for textual complexity that is plausibly uncorrelated with any unexplained component of the yield spread.¹³

The results of our IV tests are reported in Table 6. In columns (1) and (2), we test the effect of instrumented complexity on offering yield spread for high-quality bonds and medium-quality bonds, respectively, and find that the complexity-yield premia within these credit bins are close to zero. Notably, in columns (3) and (4), we find statistically significant instrumented complexity-yield premia of 18.1 basis points for low-quality bonds and 23.5 basis points for unrated bonds, respectively.¹⁴ Overall, our IV tests support the evidence from our baseline OLS tests that yields are higher when the official statement is more complex, especially for lower-quality issues.

Lastly, we consider the possibility that the bond counsel or financial advisor associated with the issue exogenously influences the complexity of that issue. In unreported tests that use bond counsel complexity as an instrument for issue complexity, we find an instrumented complexity-yield premium of 19.1 basis points for low-quality bonds, but no evidence of a complexity-yield premium for the remaining three credit categories. In additional tests that use financial advisor complexity as an instrument for issue complexity, we find an instrumented complexity-yield premium of 18.9 basis points for unrated bonds, but no evidence of a complexity-yield premium for the remaining three credit categories. These results suggest

¹³One possible reason that the underwriter instrument may not satisfy the exclusion restriction is that smaller, low-quality issuers select smaller, low-quality underwriters that tend to write complex statements. In our IV regressions, we control for the natural log of the size of the underwriter (by number of issues) to control for this channel and find similar results. Our focus on within-state average complexity and inclusion of state-year fixed effects also addresses a potential concern that state-level regulations are jointly influencing underwriter complexity and offering yields.

¹⁴If we test plain OLS regressions of offering yield spread on complexity for negotiated issues, we obtain point estimates of 9 basis points for low-quality negotiated issues and 11 basis points for unrated negotiated issues. The OLS point estimates for these last two credit bins are somewhat smaller than the IV point estimates, suggesting that the causal effect of complexity is larger than suggested by OLS regressions. One possible reason for the difference in point estimates is that the OLS regressions partially capture a component of complexity necessary for sophisticated issuers with lower default risk, even within the low-quality and unrated subcategories.

that counsel-driven complexity is especially relevant for higher-risk issuers that are more likely to face legal issues associated with potential default, and that advisor-driven complexity is especially relevant for unrated issues which tend to be more informationally opaque. However, the evidence from these tests is more suggestive because we cannot determine the extent to which the bond counsel or advisor is involved with drafting the official statement. By contrast, underwriters are more likely to assist with drafting the official statement for negotiated issues compared to competitive issues, underlining the strength of the underwriter instrument relative to the counsel or advisor instrument.

5.4 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 of 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, issue-level yield volatility has a mean (standard deviation) of 24.6 basis points (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 modest. One potential reason for the modest effect is that we lose about one-third of our issues due to a lack of secondary market trades in some issues. For example,

we find that 51% of the lost issues are unrated while only 11% of the remaining issues are unrated, implying that our reduced sample is biased toward rated issues which are less likely to be affected by complexity, particularly in the higher rating categories. Therefore, to account for the biased sample, we re-test the regression model in column (1), but also include the Complexity \times Unrated, Complexity \times Mid Rating, and Complexity \times Low Rating interactions. 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 on the near-zero premium within 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 GO bonds versus revenue bonds and high-priced bonds versus low-priced bonds, respectively. Similar to our previous results on complexity-yield premia in the cross-section, 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. Following [Green et al. \(2007\)](#), 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. We find a median roundtrip transaction cost of 90 basis points during our sample period, which closely aligns with results reported in [Griffin et al. \(2023\)](#). Issue-level liquidity is calculated 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 the remaining transactions. To maintain a balanced panel, 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 observations for the three measures. Across all issues in this analysis, the mean (standard deviation) *Markup*

is 102.9 basis points (56.1 basis points), the mean (standard deviation) *RMarkup* is 110.7 basis points (74.4 basis points), and the mean (standard deviation) *IMarkup* is 67.5 basis points (67.3 basis points). Consistent with the previous literature, the average retail markup is significantly larger than the average institutional markup.

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 is associated with reduced 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 but the size of the effect is 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 the 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 issuers could decrease their borrowing costs by simplifying their official statements, thus catering to the large population of unsophisticated investors. Despite these potential savings, complexity has been steadily increasing over time. Figure 3, Panel A indicates that complexity has increased by 0.5 standard deviations from 2010 to 2019, and Panel B indicates that residual complexity (from a regression of complexity on our bond-level control variables) has similarly increased. In terms of raw inputs, the average number of pages, number of words, and the Gunning Fog Readability grade have increased by 45 (40%), 30 thousand (64%), and 1.8 (9.0%), respectively.

Given our documented evidence of a positive and significant complexity-yield premium, why has complexity increased over time? 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 statements. This hypothesis is supported by a theoretical prediction from [Dye \(1998\)](#) that firms disclose more information (and thus increase the complexity of the disclosure) when the investor base is more informed. Further supporting this hypothesis is evidence from [Bergstresser and Cohen \(2016\)](#) of a negative trend in the percentage of households owning municipal debt, and a positive 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 catering to these investors by releasing more detailed and complex official statements. Our second hypothesis is that regulation in the municipal bond market from the SEC, MSRB, and state-level agencies has become increasingly burdensome over time, thereby increasing 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 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 percentage of sophisticated investors is increasing and municipalities are increasing 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 regulatory hypothesis, complexity trends should not differ across credit rating categories because regulatory changes implemented by the SEC or MSRB are applicable to all issuers. Consistent with the regulatory 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 percentage of sophisticated institutional investors. Therefore, we also examine trends in complexity for BQ versus non-BQ bonds and high-price versus lower-price bonds, with the intuition that the complexity trends for BQ bonds and high-price bonds will be stronger under the catering hypothesis since those bonds attract more institutional investors. In unreported tests, we find that the trends are similar across all four of these categories, suggesting that institutional catering is less likely to be the main driver of rising complexity. Lastly, we examine differences in complexity trends for bonds located in no-tax privilege states versus the remaining states. 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, suggesting 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 push to increase the number of enforcement officers in regional offices (White, 2016). Using quarterly regional SEC employment data from the Fedscope – Federal Workforce database, we find that the number of officers increased by 70% on average. There was also strong regional variation in officer growth, with the Boston office seeing an increase of 150% and the Fort 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 that experienced high enforcement officer growth relative to states that did not.

Methodologically, we define states as “high enforcement growth” (*HEG*) if they experienced above-median enforcement growth relative to other states. We find that the *HEG* SEC offices are located in Boston, Chicago, Philadelphia, San Francisco, Los Angeles, and Salt Lake City. We then regress complexity on indicator variables for each year, interactions of each year with *HEG*, 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.¹⁵ Therefore, our evidence suggests that regulatory enforcement is an important driver of complexity.

Lastly, we decompose complexity into four categories to test if the complexity-yield premium loads heavily on any particular category. We first use the Latent Dirichlet Allocation (LDA) model (Blei et al., 2003) to identify LDA topics discussed within the text of each official statement. LDA is an unsupervised machine learning model that represents each individual document 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 (the full sample of official statements), the construction of topics takes place without decisions regarding the subject matter. This technique removes the potential for biases that could arise from a list of predetermined keywords, for example. While the model determines the composition of words for each topic, we assign each topic to a broader category in order to facilitate interpretation.

The LDA approach is especially useful for our analysis because the municipal bond official statements 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, we would expect topic weightings to vary substantially within the document. Therefore, we estimate the LDA model at the page level for the purpose of recovering broad topic weights for each page. We first use the LDA model to group textual information into 100 LDA topics for each page of each official statement.¹⁶ We assign the majority of LDA topics to one of four broad categories: (1) risk, (2) bond description, (3) cash flow, and (4) legal. For some LDA topics, the category is not clear, and those topics remain unclassified. In Table 9, we provide the top three word

¹⁵This result is consistent with findings in Maffett et al. (2023) that large underwriters were more likely to enhance disclosure in the municipal bond offering market following the introduction of the SEC's Municipalities Continuing Disclosure Cooperation (MCDC) initiative in 2014. These large underwriters are more likely to be located in large cities with HEG SEC offices, and the enhanced disclosure was also likely associated with an increase in official statement length and complexity.

¹⁶An LDA topic is the dot product of a vector of weights and a vector of words (e.g., topic $X = 0.4 \times \text{"entity"} + 0.3 \times \text{"budget"} + 0.2 \times \text{"fund"} + \dots$). Each page of an official statement is comprised of the weights assigned to each LDA topic.

loadings for each of 12 select LDA topics identified by the model, and the mappings of these topics to the four broad categories. For example, we assign the LDA topic of “Security, Risk, Fund” to risk, “Project, Improvement, Capital” to bond description, “Expense, Fee, Cost” to cash flow, and “Bond, Opinion, Counsel” to legal.

For each official statement, we calculate four complexity measures based on these four broad categories using the following methodology. First, for each page, we compute the product of each complexity input (such as the Gunning Fog readability score) and each of the four broad category weights. For example, if page 15 of a particular official statement has a Gunning Fog readability score of 15 and a risk weight of 0.30, then the risk category-weighted Gunning Fog readability score is calculated as $15 \times 0.30 = 4.5$. Second, we aggregate each category-weighted complexity input to the official statement level by taking its average across all pages in the official statement. Third, we standardize each aggregate measure to have a mean of zero and a standard deviation of one using our full sample of official statements. Following this third step, an official statement could have a standardized risk category-weighted Gunning Fog readability score of 0.5, a standardized risk category-weighted word count of 0.8, and so on. Fourth, for each official statement and category, we calculate the average standardized complexity measure to obtain four complexity metrics for each official statement based on risk, bond description, cash flow, and legal. The third and fourth steps are similar to how we calculate our baseline complexity measure, except in this case we are calculating four complexity measures based on the four categories.

We re-test the regression model in equation (12) for each of our four credit rating bins, except that we replace our broad complexity measure with our four categorical complexity measures. The results are reported in Table 10. In columns (1) and (2), we find that legal complexity is the main driver of yield for high-quality and medium-quality bonds, with a 4.3 basis point and 3.6 basis point increase in yield for every standard deviation increase in legal complexity, respectively. In both of these columns, description complexity also has a small effect. In columns (3) and (4), we again find that legal complexity is the main driver of yield for low-quality and unrated bonds (12.4 basis points and 19.7 basis points, respectively), and that the economic magnitudes of the legal complexity effects are much stronger compared to the higher-quality issues. For low-quality and unrated bonds, we find that risk complexity and description complexity also 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.

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 exogenously 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 complexity in each statement, and provide evidence of a large complexity-yield premium. Importantly, the size of this premium very much depends on the clientele base. When sophisticated institutional investors comprise a larger percentage of the clientele, which is the case for bonds that are bank-qualified or issued in states without local tax privileges, then we find a near-zero premium. Conversely, the complexity-yield premium is large when the clientele base is comprised of 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 for smaller issuers. In supporting spatial IV tests that use average underwriter complexity as an instrument for issue complexity, we also find evidence of a sizable complexity-yield premium, suggesting a causal interpretation of our central results.

Despite the significant complexity-yield premium, official statement complexity has been steadily rising over time. We test two hypotheses to uncover the reasons for this puzzling trend: (1) issuers are increasingly catering to the growing percentage of sophisticated investors ([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 reg-

ulatory hypothesis. When we decompose complexity into multiple categories 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 smaller issuers that rely on local investors or lack coverage from a credit rating agency. One possible solution for reducing the costs associated with this premium is to require that issuers provide a brief executive summary of relevant risks at the beginning of the official statement so that resource-constrained investors do not necessarily have to pore through long, complicated statements to identify those risks. For smaller municipalities that are less likely to obtain a credit rating due to budget constraints, an executive summary that outlines the key risks would be especially effective in reducing their borrowing costs and promoting economic growth.

References

- Acemoglu, D., Johnson, S., Robinson, J. A., Yared, P., 2008. Income and democracy. *The American Economic Review* 98, 808–842.
- 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.
- Aggarwal, Chaitanya; Bansal, S., 2023. Textstat. <https://github.com/textstat/textstat>.
- 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 Rule-making 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.
- Butler, A. W., Yi, H., 2022. Aging and public financing costs: Evidence from US municipal bond markets. *Journal of Public Economics* 211, 1–18.

- 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.
- Dye, R. A., 1998. Investor sophistication and voluntary disclosures. *Review of Accounting Studies* 3, 261–287.
- 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.

- Giesecke, O., Mateen, H., Sena, M., 2023. Local government valuation. Working paper. Stanford University.
- Goldsmith-Pinkham, P., Gustafson, M. T., Lewis, R. C., Schwert, M., 2023. Sea-level rise exposure and municipal bond yields. *Review of Financial Studies* 36, 4588–4635.
- 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.
- 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.
- 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., 2017. The use of EDGAR filings by investors. *Journal of Behavioral Finance* 18, 231–248.
- Loughran, T., McDonald, B., 2020. Textual analysis in finance. *Annual Review of Financial Economics* 12, 357–375.
- Maffett, M., Samuels, D., Zhou, F., 2023. Carrot or stick? The impact of regulatory leniency on municipal disclosure compliance. Working paper. University of Miami.
- 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.
- Municipal Bonds for America, 2016. What are muni bonds? <https://munibondsforamerica.org/resources/what-are-muni-bonds/>.
- NerdWallet, 2020. Humans vs. robots: Americans prefer financial advisors over algorithms. <https://www.nerdwallet.com/article/investing/robo-advisor-survey>.

- Northwestern Mutual, 2022. The majority of Americans say their financial planning needs improvement, but only a third seek professional help. <https://news.northwesternmutual.com/2022-07-27-The-Majority-of-Americans-Say-Their-Financial-Planning-Needs-Improvement,-But-Only-a-Third-Seek-Professional-Help>.
- Novy-Marx, R., Rauh, J. D., 2012. Fiscal imbalances and borrowing costs: Evidence from state investment losses. *American Economic Journal: Economic Policy* 4, 182–213.
- Painter, M., 2020. An inconvenient cost: The effects of climate change on municipal bonds. *Journal of Financial Economics* 135, 468–482.
- Persson, T., Tabellini, G., 2009. Democratic capital: The nexus of political and economic change. *American Economic Journal: Macroeconomics* 1, 88–126.
- 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>.
- WM Financial Strategies, 2013. Bond sale methods (competitive vs. negotiated bond sales). <https://www.munibondadvisor.com/SaleChoice.htm>.

- 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 SE-
RIES 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

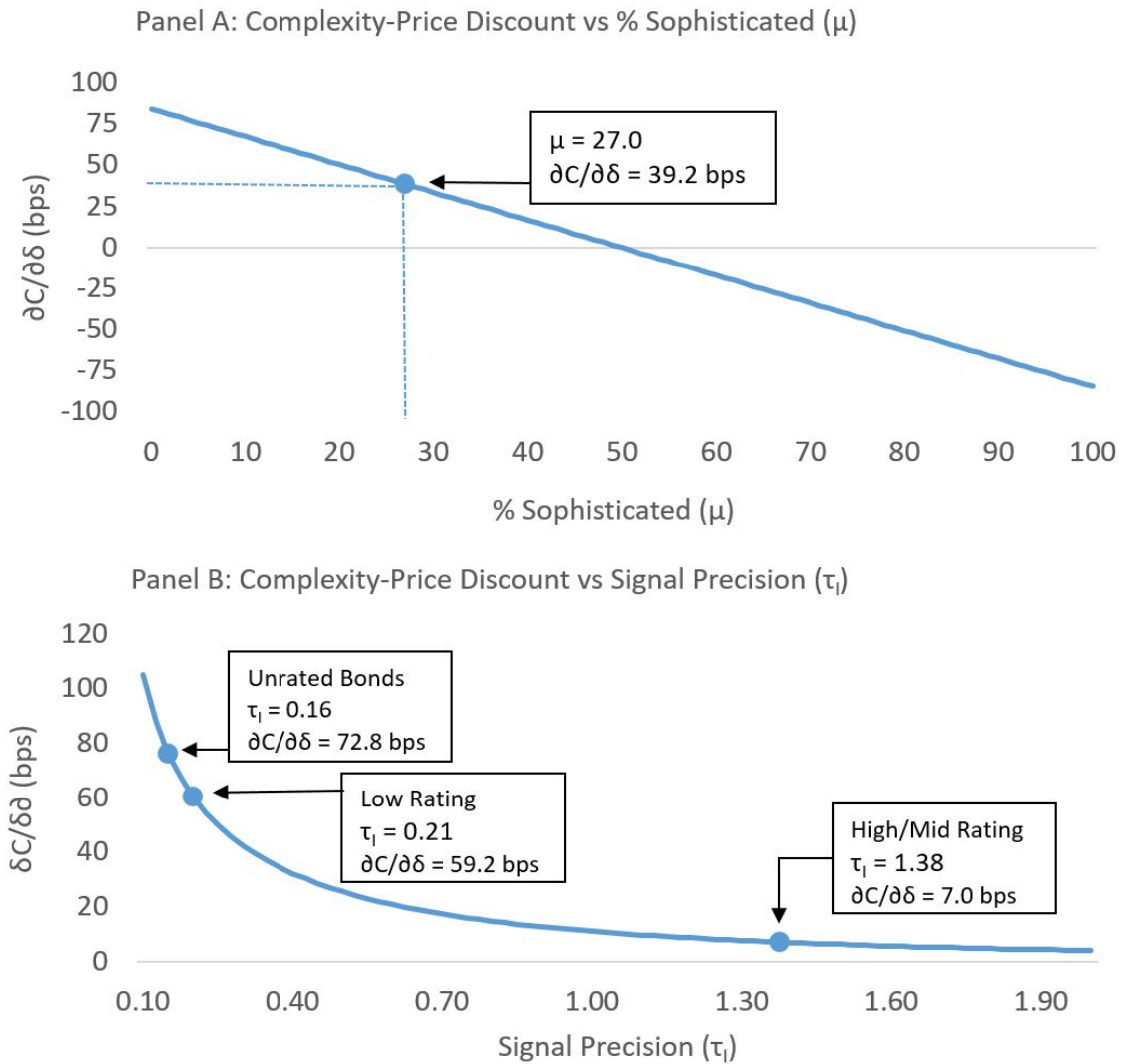
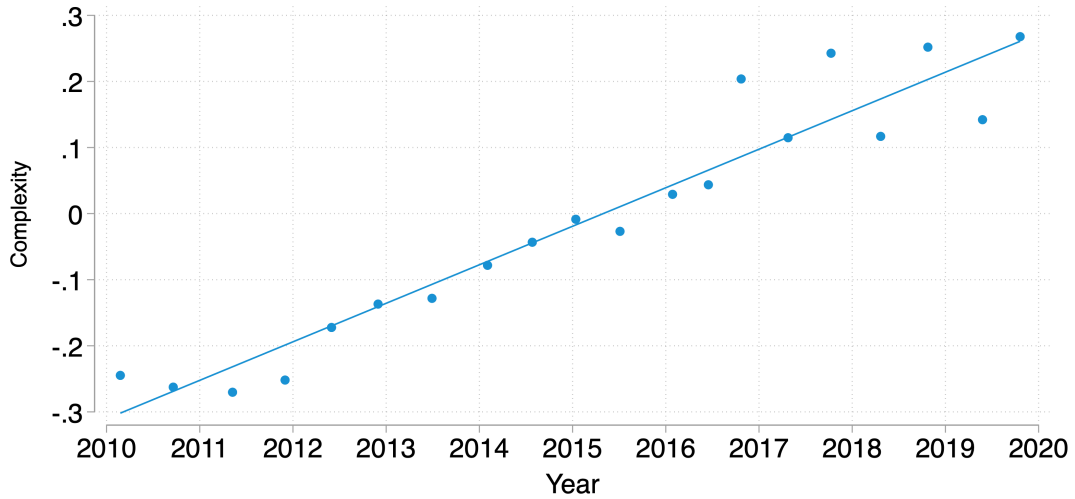


Figure 2: Complexity-Price Discount versus Clientele Base and Signal Precision. Panel A graphs the complexity-price discount versus the proportion of sophisticated traders (μ). Complexity-price discount is calculated as $(\partial C/\partial \delta) \times (0.35/100)$, where the expression for $\partial C/\partial \delta$ 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. Complexity-price discount 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 complexity-price discount 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 Complexity Over Time



Panel B: Residualized Complexity Over Time

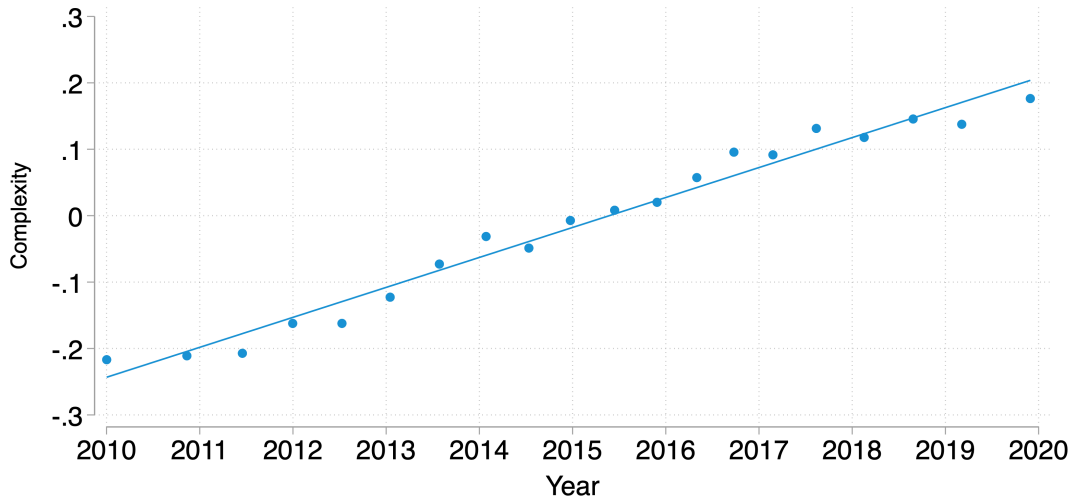


Figure 3: Time Series of Complexity in Official Statements. Panel A reports average complexity for each year during our sample period of mid-2009 to mid-2020. Panel B reports average residual complexity for each year during our sample period, where residual complexity is obtained from a regression of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

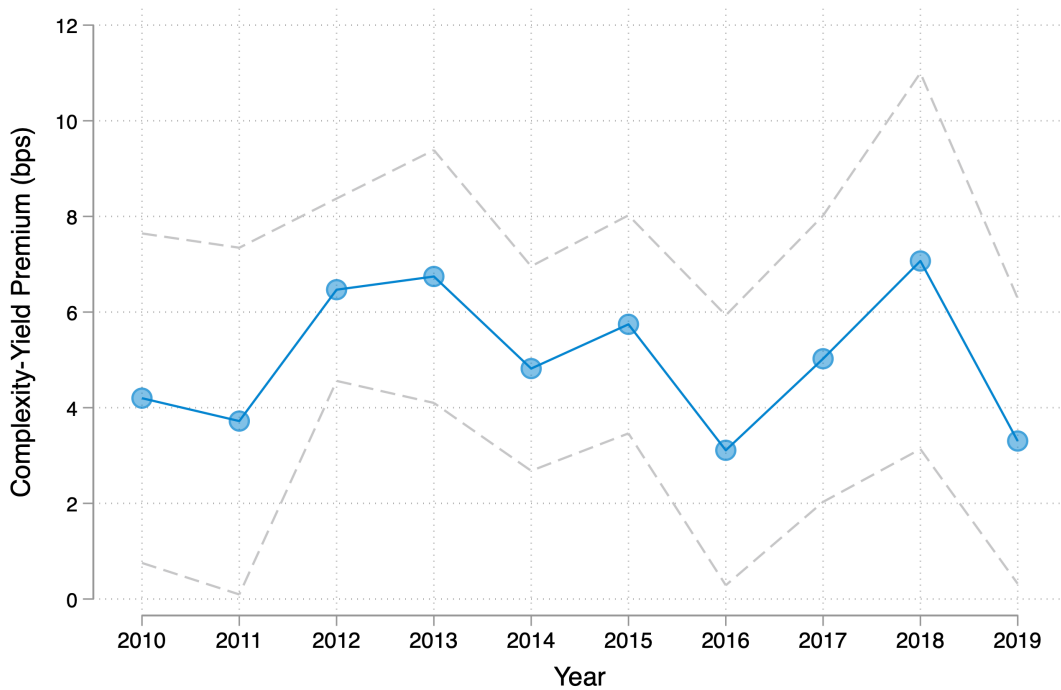


Figure 4: Complexity-Yield Premium by Year. This graph plots the complexity-yield premium for each full year of our sample period. The complexity-yield premium for each year is obtained by estimating our baseline regression model with the inclusion of interactions of complexity and an indicator variable 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.

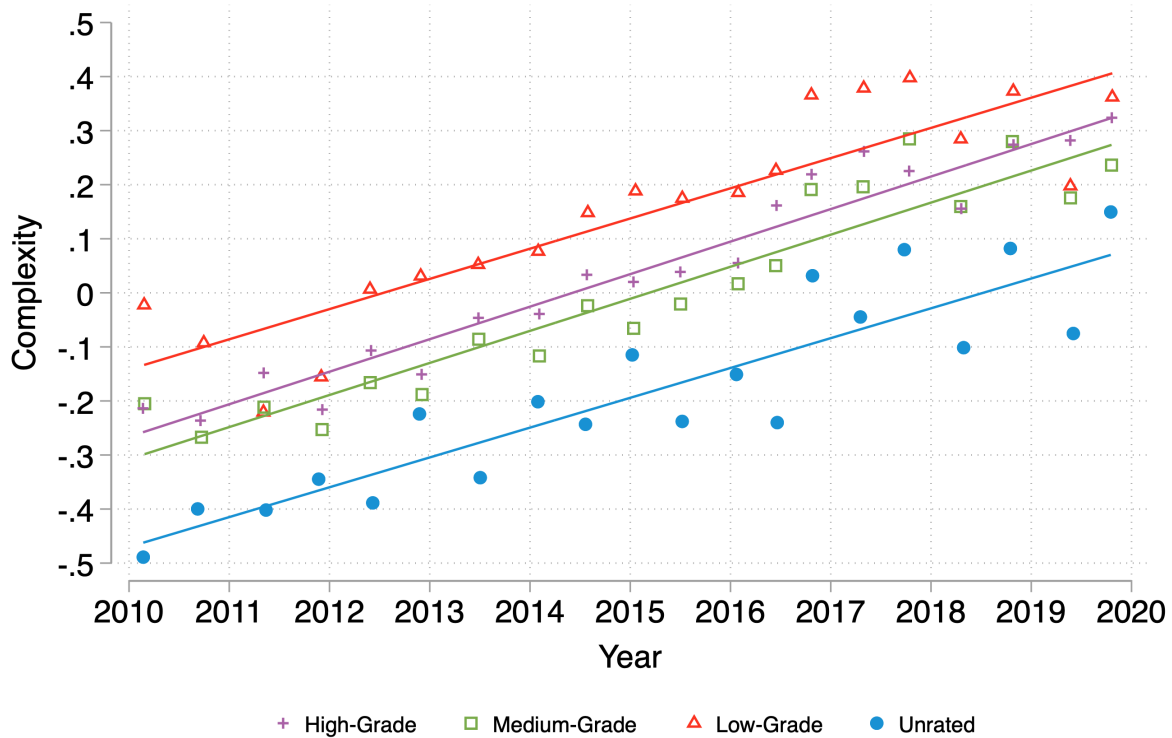


Figure 5: Complexity by Rating Category. This figure presents linear trends in municipal bond official statement complexity during our sample period for the following rating categories: high-rated bonds, medium-rated bonds, low-rated bonds, and unrated bonds.

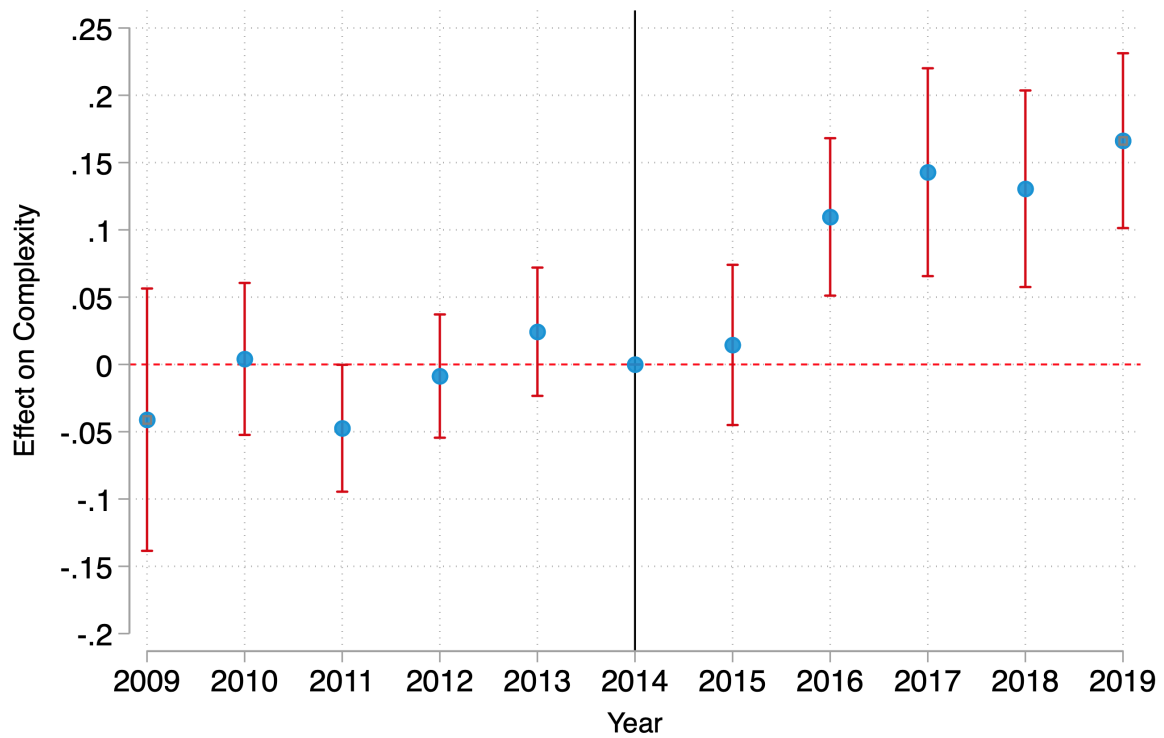


Figure 6: Effect of Regulatory Burden on Official Statement Complexity. This figure shows the evolution of complexity for states that experience above-median growth in SEC enforcement officers (*HEG* states) relative to states that experience below-median growth during the SEC enforcement push from 2015 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. Columns (3) also includes state-year fixed effects, and column (4) includes state-year (SY) and issuer fixed effects. 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*** (0.013)	0.034*** (0.009)	0.049*** (0.008)	0.023*** (0.005)
Issue Size		-0.059*** (0.007)	-0.054*** (0.007)	0.006 (0.004)
Maturity		0.187*** (0.021)	0.192*** (0.023)	0.110*** (0.013)
Negotiated		0.110*** (0.016)	0.118*** (0.016)	-0.019* (0.011)
Gen. Obligation		-0.298*** (0.015)	-0.295*** (0.015)	-0.029*** (0.009)
Insured		0.037** (0.016)	0.033** (0.015)	-0.116*** (0.021)
Callable		0.061*** (0.017)	0.050*** (0.016)	0.062*** (0.009)
ln(Population)		0.037*** (0.007)	0.027*** (0.006)	-0.159 (0.127)
Unemp. Rate		0.054*** (0.007)	0.020*** (0.003)	0.007 (0.008)
Fixed Effects	None	None	State-Year	SY, Issuer
R^2	0.022	0.243	0.346	0.735
N	84,546	84,545	84,542	72,454

Table 3: Complexity and Yield Spread in the Cross-Section. Columns (1) to (4) of Panel A report the results of OLS regressions of municipal bond yield spread on complexity for the subsamples of high-rated, medium-rated, low-rated, and unrated issues, respectively. Columns (1) to (4) of Panel B report the results of OLS regressions of municipal bond yield spread on complexity for GO issues, high/medium-rated revenue issues, low-rated revenue issues, and unrated revenue issues, respectively. The “High” and “Medium” rating subsamples represent the issues that are placed in the highest two credit rating categories and next-highest two credit rating categories, respectively. The “Low” rating subsample represents the issues that are placed in all other credit rating categories. 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.

Panel A: Complexity-Yield Premium by Credit Rating Category				
	(1)	(2)	(3)	(4)
	Yield Spread	Yield Spread	Yield Spread	Yield Spread
Complexity	0.013*** (0.004)	0.006* (0.003)	0.074*** (0.008)	0.091*** (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

Panel B: Complexity-Yield Premium by GO/Revenue Category				
	(1)	(2)	(3)	(4)
	Yield Spread	Yield Spread	Yield Spread	Yield Spread
Complexity	-0.024*** (0.005)	0.037*** (0.005)	0.105*** (0.011)	0.140*** (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 4: 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) reports the results from a similar test that instead uses the remaining states with tax privileges. In column (3), we utilize the full sample and include 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 utilize the full sample and include 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 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)	(2)	(3)	(4)
	Yield Spread	Yield Spread	Yield Spread	Yield Spread
Complexity	0.023 (0.015)	0.056*** (0.010)	0.061*** (0.010)	0.067*** (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 5: 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 that equals one if the bond issuance price is at least \$103. In column (2), complexity is interacted with a “BQ” indicator that equals one if the bond is bank qualified. In column (3), complexity is interacted with a “Negotiated” indicator that 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)	(2)	(3)	(4)
	Yield Spread	Yield Spread	Yield Spread	Yield Spread
Complexity	0.066*** (0.011)	0.090*** (0.013)	0.032*** (0.009)	0.105*** (0.017)
Complexity × BQ	-0.067*** (0.013)			-0.083*** (0.014)
Complexity × High Price		-0.096*** (0.013)		-0.113*** (0.014)
Complexity × Negotiated			0.044*** (0.014)	0.043*** (0.012)
Controls	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.357	0.395	0.346	0.409
N	84,542	70,902	84,542	70,902

Table 6: Spatial IV Tests of the Complexity-Yield Premium. Each numbered column reports the results of a spatial IV regression of negotiated municipal bond yield spread on instrumented complexity within a particular rating category. The instrument for complexity is the average within-state complexity of the underwriter involved with the negotiated issue, where the average excludes any observations associated with that issuer. 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. First-stage Kleibergen-Paap (KP) F statistics are reported in the last row of the table. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	Yield Spread	Yield Spread	Yield Spread	Yield Spread
Pred. Complexity	0.020 (0.068)	0.021 (0.065)	0.181** (0.089)	0.235*** (0.080)
Rating	High	Medium	Low	Unrated
Controls	Yes	Yes	Yes	Yes
State-Year FE	State-Year	State-Year	State-Year	State-Year
R^2	0.201	0.248	0.259	0.374
N	4,098	9,116	7,011	3,885
KP- F	8.94	32.95	25.51	81.22

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 or more 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 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)	(2)	(3)	(4)
	Yield Vol.	Yield Vol.	Yield Vol.	Yield Vol.
Complexity	0.003*** (0.001)	-0.002 (0.001)	0.006*** (0.002)	0.008*** (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 that we use issue-level average retail roundtrip transaction cost (RMarkup) and issue-level average institutional roundtrip transaction cost (Inst. Markup) as the dependent variables, 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)	(2)	(3)
	Markup	Ret. Markup	Inst. Markup
Complexity	0.012*** (0.004)	0.017*** (0.005)	-0.011** (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 word loadings for each of 12 select LDA-generated topics. We assign these LDA-generated topics to one of the following four categories: Risk, Bond Description, Cash Flow, and Legal.

Risk	Bond Description
<ol style="list-style-type: none"> 1. Security, Risk, Fund 2. Liability, Deferred, Resource 3. Default, Event, Trustee 	<ol style="list-style-type: none"> 1. Project, Improvement, Capital 2. Redemption, Bond, Date 3. Bond, Price, Discount
Cash Flow	Legal
<ol style="list-style-type: none"> 1. Expense, Fee, Cost 2. Budget, Appropriation, Year 3. Fund, Investment, Portfolio 	<ol style="list-style-type: none"> 1. Bond, Opinion, Counsel 2. Mean, Resolution, Shall 3. Authority, Finance, Resolution

Table 10: Categorical Complexity and Yield Spread. Each column reports the results of an OLS regression of municipal bond yield spread on the following four categorical complexity measures: risk complexity, legal complexity, cash flow complexity, and description complexity. 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)	(2)	(3)	(4)
	Yield Spread	Yield Spread	Yield Spread	Yield Spread
Risk Complexity	0.003 (0.005)	-0.003 (0.005)	0.046*** (0.010)	0.040*** (0.013)
Legal Complexity	0.043*** (0.006)	0.036*** (0.005)	0.124*** (0.011)	0.197*** (0.021)
Cash Flow Complexity	0.003 (0.005)	0.005 (0.005)	0.002 (0.008)	0.007 (0.009)
Description Complexity	0.013** (0.005)	0.016*** (0.004)	0.026*** (0.008)	0.046*** (0.012)
Rating	High	Medium	Low	Unrated
Controls	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.375	0.475	0.504	0.650
N	17,886	28,154	16,536	20,387