Textual Disclosure in Prospectuses and Investors' Security Pricing *

Jörn Debener † Arved Fenner ‡ Philipp Klein § Steven Ongena ¶

October 13, 2023

*We thank Konrad Adler, Christoph Basten, Tobias Berg, Winta Beyene, Nicole Branger, Denis Davydov, Günter Franke, Monika Gehde-Trapp, Michael Goedde-Menke, Martin Hibbeln, Maximilian Jager, Patrick Kampkötter, Florian Kiesel, Michael Koetter, Christian Leuz, Friedrich Lorenz, Ouarda Merrouche, Huyen Nguyen, Andreas Pfingsten, Max Riedel, Christoph Schneider, Salomon Garcia Villegas, Alexander Wagner, and the participants at the International Finance and Banking Society Oxford Conference in 2021, the International Risk Management Conference in 2021, the Research Seminar in Contract Theory, Banking and Money at the University of Zürich in 2021, the Münster Banking Workshop in 2021, the Australasian Finance and Banking Conference in 2021, the Finance Center Münster Research Seminar in 2021, the 38th International Conference of the French Finance Association in 2022, the Business Research Seminar at the University of Tübingen in 2022, the Summer Meeting of the International Banking, Economics and Finance Association in 2022, the Annual Meeting of the European Financial Management Association in 2022, the European Meeting of the Financial Management Association in 2022, the Annual Meeting of the European Economic Association in 2023, the Annual Conference of the German Society for Operations Research in 2023, the GRETA C.R.E.D.I.T. conference in 2023, and the Annual Meeting of the German Finance Association in 2023 for providing us with very helpful suggestions. Philipp Klein gratefully acknowledges financial support from the German Research Foundation (Deutsche Forschungsgemeinschaft) – grant no. 466168740.

†Jörn Debener, University of Münster, Email: joern.debener@wiwi.uni-muenster.de.

‡Arved Fenner, University of Münster, Email: arved.fenner@wiwi.uni-muenster.de.

§Philipp Klein, University of Münster, University of Zürich, Email: philipp.klein@wiwi.uni-muenster.de.

¶Steven Ongena, University of Zürich, SFI, KU Leuven, NTNU Business School, and CEPR, Email: steven.ongena@bf.uzh.ch.

Abstract

We explore the impact of textual disclosures' quality and quantity, measured as the share of boilerplate language, the linguistic complexity, and the disclosure length, on investors' security pricing at issuance. Exploiting an extensive data set covering over 1,000 issuance prospectuses of ABS transactions involving almost 40 million loans, we show that the prospectuses' quality and quantity substantially affect investors' pricing beyond all observable risk factors. Investors demand an economically significant higher yield spread if the share of boilerplate language decreases or if prospectuses are lengthy. Hence, more content, in relative or absolute terms, scares investors. To explain these surprising findings, we investigate the importance of three possible mechanisms: presumed default risk, level of information asymmetry, and visualizations supplementing the prospectus. We also document that a lower prospectuses' quality and a higher quantity weaken investors' risk assessment, decreasing the economic efficiency in the bond market. Recent EU regulations aiming at addressing these problems have homogenized the quality and quantity of textual disclosure in ABS prospectuses. Our results have important implications for market participants and regulators alike, placing the quality and quantity of textual disclosure in prospectuses high on their agenda.

Keywords: Textual Disclosure, Prospectuses, ABS, Textual Analysis
JEL Classification: G11, G14, G21, G23

I Introduction

For an adequate pricing of securities, understanding their risks is crucial. For this purpose, comprehensive information is necessary to enable investors to conduct an appropriate risk assessment. Issuance prospectuses contain important information on the risks and potential returns of newly issued securities. Therefore, these prospectuses play a crucial role in the investment decisions of investors.[1] This is particularly pronounced for the issuance of asset-backed securities (ABS), where investors—typically institutional investors like mutual funds or insurance companies—have limited publicly available information on essential security characteristics like the underlying loan portfolio and the deal structure and, therefore, have to rely on the information communicated in the issuance process. The institutional framework in our study allows us to overcome the regularly raised concern in the literature on investors' access to public information potentially processed in their investment decisions (Bourveau et al., 2022).

Due to the central role of prospectuses in the issuance process, regulators put a lot of effort into ensuring that prospectuses are as informative as possible and serve as a means of investor protection (see, for example, the EU Prospectus Regulation in 2017 or, for the US, the Securities Offering Reform in 2005). However, the European Securities and Markets Authority is still concerned in 2022 that issuance prospectuses may "contain substantial repetition of text, [...], may present generic and imprecise risk factors, and may include unclear language" and thus "an abundance of material can present a challenge for even specialized readers to identify information that is key to assessing the product" (European Securities and Markets Authority, 2022).

Issuers have different incentives, like avoiding high yield spreads demanded by investors, having low transaction costs associated with the issuance process, and reducing legal risks over the security lifetime, and may design prospectuses accordingly (Hanley and Hoberg, 2010; Howell, 2018). Therefore, how the information presented in prospectuses affects

[1] We use the terms issuance prospectus and prospectus interchangeably.

investors is unclear and requires empirical investigation. Although a small but growing literature indeed aims to obtain a deeper understanding of the prospectuses' effect on investors (e.g., Abis et al., 2021; deHaan et al., 2021; Zhang et al., 2023), the question remains how the quality and quantity of textual disclosure in prospectuses affect investors in their security pricing.

We exploit a novel and extensive data set comprising detailed loan, security, and deal information, as well as prospectuses of more than 1,000 European ABS deals between 2002 and 2020. Our data set includes over 39.4 million individual loans representing all deals reported to the European DataWarehouse (EDW), the only central securitization loan-level repository under the loan-level initiative (LLI) of the European Central Bank (ECB). We enrich this data set with pricing and rating data from S&P Global and interest rate spreads from the ECB, Refinitiv Datastream, and Federal Reserve Economic Data (FRED). Based on this unique data set, we measure the share of standard phrases, commonly referred to as boilerplate language, the linguistic complexity, and the length of ABS prospectuses.

We find that a higher share of boilerplate language in ABS prospectuses is associated with significantly lower initial yield spreads, indicating that investors interpret standardized descriptions of ABS characteristics as a positive signal for the securities' risk-return profile. Specifically, an increase of the share of boilerplate language by one standard deviation lowers the demanded yield spreads by 17 basis points (bp) representing about 18% of the mean initial yield spread of 91 bp in our sample. Lengthier ABS prospectuses, however, are associated with higher yields, suggesting that investors perceive lengthy prospectuses as a warning signal of a higher default risk. Here, an increase by one standard deviation leads to an increase of the initial yield spread by 12 bp representing more than 13% of the mean yield spread. Thus, our results show that the use of boilerplate language but short prospectus descriptions let investors assume lower risks over the security lifetime, while more definite and longer descriptions make investors wary and lead them to surmise additional, potentially unobserved risks. To explain these surprising findings, we examine

three potential mechanisms of how textual information in the prospectuses affects the investors' security pricing behavior.

First, we take advantage of the large securities' default risk variation due to the subordination of repayments as a unique setting of ABS transactions in our data set. The greater use of boilerplate language leads to lower yield spreads and thus is positively perceived by investors if the investors expect a low level of security's default risk as shown by a high initial credit rating, which will only materialize in strong economic downturns, and reverses with higher investors' expectation of losses. In line, investors' demand for higher yield spreads in case of particularly lengthy and less concise prospectuses only applies to securities with a substantial default risk.

Second, we exploit that the issuing banks in our sample significantly differ in their securitization activity, and thus, as investors may learn from previous deals, in the level of asymmetric information between banks and investors. Our results show that the decreasing impact of boilerplate language on the yield spread is particularly significant if the level of information asymmetries is high, and thus the investors more strongly rely on the issuance prospectuses, and reduces as the information asymmetry declines. Moreover, lengthier prospectuses have a particularly strong raising effect on the yield spread when information asymmetry is low, and thus, investors seem to be negatively surprised when the bank unusually provides exceptionally long descriptions of one of its transactions and fear abnormally great risks. This finding underpins the interpretation that a high share of boilerplate language and a concise prospectus are perceived as a signal of less conspicuous risks, which investors may be particularly concerned about if they have less previous experience in dealing with the bank.

Third, visualizations make information easier to comprehend and investors' work more efficient, but they also curtail complex content and reduce the comparability of information (e.g., Lurie and Mason, 2007). Using the file size as a measure of the visualization extent in the highly individualized prospectuses in our sample, we find that visualization and standardized language behave as complements, and thus, a higher level of visualization

amplifies the decreasing impact of boilerplate language on the yield spread. Additionally, visualization weakens the increasing impact of the prospectus length on the yield spread.

The texts in our sample's prospectuses are exceedingly complexly written as they exhibit an average Gunning's FOG index of almost 24,[2] while the average FOG index of articles in the Wall Street Journal is between 11 and 12, of a typical dissertation in the US between 16 and 17 (Cox, 2007), and, exemplarily, of initial coin offerings (ICO) white papers in the study by Bourveau et al. (2022), it is 15.9. As presumed readers, there are professional, institutional investors, primarily banks, investment funds, and insurance companies, while it is widely prohibited to sell ABS to retail customers in Europe (Regulation (EU) 2017/2402, European Systemic Risk Board (2022)). However, the majority of these investors only hold an undergraduate degree (at least in the US, Zippia, Inc. (2023)), potentially raising doubts as to whether the prospectuses are fully understood by investors and further highlighting the relevance of the textual design in issuance prospectuses for their pricing decisions.

In line, we provide evidence that more boilerplate language, more complex descriptions, and a higher textual disclosure quantity decrease the predictive ability of initial yield spreads demanded by investors for future security performance. This indicates that investors' risk assessment is less accurate and the efficient allocation of risk premiums to risky securities is distorted. We also show that investors, as a consequence of the less accurate pricing at the security issuance, adjust their pricing during the security lifetime more strongly for securities with more complex and long prospectuses. Our results highlight the importance of textual information disclosure in issuance prospectuses for the economic efficiency in the bond market in particular and of financial markets in general.

Our findings do not result from correlations of textual disclosure quality and quantity with initial loan portfolio risk or subsequent security performance, nor with the complexity of the underlying ABS deal structure or the underlying collateral. Furthermore, we find

[2]The Gunning's Fog index is a commonly used measure of linguistic complexity in the literature on financial disclosures. The index is initially designed with the broader interpretation signaling the number of years of formal education that is necessary for an average reader to understand the text after the first reading (e.g., Li, 2008). The FOG index of our paper is 21.65 (based on the version from March 9, 2023).

the same relationship between textual disclosure quality and quantity on the one hand and investors' pricing on the other hand for the risk factor section, which is arguably very important for investors' security pricing, and when additionally controlling for the involved law firms assisting the bank by preparing the prospectus. Instead, the use of boilerplate language and linguistic complexity is significantly driven by low banks' regulatory capital and high non-performing loan (NPL) ratios, while the bank size and a high liquid asset ratio affect the prospectus length. This further confirms that our results are not driven by the issuing banks' characteristics or those of the underlying loan portfolio , which is again reinforced when including the bank characteristics as controls in our analyses. Finally, we find empirical evidence based on a difference-in-differences approach proposed by deHaan et al. (2021) that two recent EU regulations have homogenized the textual disclosure quality and quantity in ABS prospectuses. Specifically, issuers with particularly standardized, complex, and lengthy prospectuses before the regulations have predominantly adjusted their prospectus design.

We contribute to the literature in several ways. First, we extend the discussion on the impact of information provision for investors' pricing of securities and ABS in particular and the adequacy of the resulting initially demanded yield spread. Besides the findings from Zhang et al. (2023) that written communication contains important information that is not fully priced by many investors, Neilson et al. (2022) show that more granular and easier accessible quantitative information on the underlying loans in ABS enhances the accuracy of investors' pricing. Our results add to these findings by emphasizing the importance of the quality and quantity of the disclosed textual information in investors' risk assessment and pricing decisions.

More broadly, our study contributes to the literature on contractual and textual information processing in financial markets (e.g., Blankespoor et al., 2020; Bourveau et al., 2022; Umar, 2022; Chy and Kyung, 2023). Bourveau et al. (2022) analyze ICO white papers, a widely individual and unregulated market, in which the documents need only 5000 to 6000 words, and find that the investors reward more and lengthier disclosures. In contrast,

the issuance prospectuses in our sample have an average length of about 92,000 words, highlighting the provision of extended explanations and, likewise, the banks' effort to secure themselves from legal risks. For this similar but stricter regulated market conditions, our study shows that the textual explanation amount is negatively perceived by investors if these assume significant default risks and have recourse to historical data, indicating that the generalized impact of the prospectus length on the investors' pricing and investment behavior is non-monotonic.

Second, we contribute to the discussion on the utilization of boilerplate language and standardized formulations and its impact on economic mechanisms and the addressees' perception of risk. A number of studies provide evidence that parties being especially concerned about their reputation are incentivized to use standardized language for precautionary purposes, and that regulators and courts react positively to it (e.g., Meade and Stasavage, 2008; Hansen et al., 2017; Cazier et al., 2021; Kleymenova and Tomy, 2022). In line, the frequency of corporate disclosures is exposed to the peer effect, particularly pronounced when strategic uncertainty and reliance on external funding are high (Seo, 2021). While investors, however, are getting used to standardized disclosures by central banks and deviations from this lead to strongly increased market volatility (Ehrmann and Talmi, 2020), the results on corporate disclosures in the literature so far show that investors, analysts, and other market participants react negatively to boilerplate language as market liquidity, institutional ownership, and analyst coverage are found to be lower and risk premiums and market volatility are higher when firms use more boilerplate language in their communications (Lang and Stice-Lawrence, 2015; McClane, 2019). Our study provides novel insights into the effect of boilerplate language on investors in financial markets by ascertaining that boilerplate language is not inevitably negatively perceived by investors if the level of information asymmetries between issuers and investors is high and the market structure is opaque and individualized.

Finally, by examining textual disclosure in prospectuses, we add to the literature analyzing textual characteristics in corporate disclosures and their influence on the economic behavior

of companies and investors, as well as the resulting market outcomes (e.g., Li, 2008; Kravet and Muslu, 2013; Lawrence, 2013; Lang and Stice-Lawrence, 2015; Dyer et al., 2017; Lo et al., 2017; Bushee et al., 2018; Cazier et al., 2021). While the subjects to be covered in the prospectus are largely specified by regulation in our sample, the way the information is presented and its textual phrasing is multifarious. Our results complement the existing literature by showing that linguistic complexity does not affect the average level of the initially demanded yield under the circumstances in our sample. As one of the most important characteristics, there are only institutional and experienced investors dealing with large investment amounts in the market. However, even for those investors, the ability to adequately allocate risk premiums to actually risky securities decreases, leading to higher uncertainty and lower economic efficiency. Our results also add to the discussion on the impact of the prospectus' design on investors, strongly shifted into the focus of the scientific debate by Celérier and Vallée (2017). While their results apply to retail customers, our study provides insights into the institutional investor side of the market.

Our results are important for market participants and regulators alike. Issuers should avoid overly long prospectuses to prevent higher demanded initial yield spreads by investors. Investors, in turn, should try to look behind very standardized prospectus designs when pricing securities. Regulators should continue to pay attention to the textual design of prospectuses to prevent undesired consequences for the market efficiency and stability.

The remainder of the paper is structured as follows. Section II provides an overview of the regulation and research on prospectuses in general and ABS prospectuses in particular. Section III presents our data sources and the sample selection process. Section IV describes the construction of the main variables, including measures of textual disclosure quality and quantity, and provides descriptive statistics. Section V presents the analysis regarding the effects of textual disclosure quality and quantity in ABS prospectuses on investors. Section VI concludes.

II Regulation and research on prospectuses

II.1 Regulatory background

When securities are issued, information on these securities is asymmetrically distributed between the issuer and potential investors (Myers and Majluf, 1984). To warrant investor protection and to enhance capital market efficiency, regulators have a vital interest in reducing asymmetric information during the security issuance process (Howell, 2018). Prospectuses are the central instrument in this task because they typically contain extensive and detailed descriptions of all important information on the issued securities and their issuer. Besides, for some securities, there are also insights into loan-level data or, more frequently, a limited number of smaller documents from other parties involved in the issuance process, such as rating agencies, available, providing summarized information (i.e., new issue reports or pre-sale reports). However, the prospectus remains the most complete and detailed source of information for investors and constitutes the legal agreement between issuers and investors.

The relevance of prospectuses from the perspective of regulators is reflected in several regulations, which specify some conditions for the information that prospectuses have to contain. In this way, these regulations provide additional motivation for our study and represent the institutional framework of our empirical analysis. Therefore, we briefly describe some of the most important regulations in the following. Because the prospectuses examined in our study originate from Europe between 2002 and 2020 (see Section III for more information on our sample), we focus on the most important regulations in the European Union during this period. Note, however, that prospectuses are also highly regulated in other jurisdictions.[3]

As the first significant regulation in this period, the European Union adopted Directive 2003/71/EC of the European Parliament and of the Council ("EU Prospectus Directive")

[3]See, for example, the Securities Offering Reform in 2005 by the US Securities and Exchange Commission.

in 2003. This directive applied to all securities offered to the public (Article 3(1)) or admitted to trading on a regulated market in the European Union (Article 3(3)).[4] The aim of this directive was to strengthen the European financial market by harmonizing the requirements for prospectuses while at the same time ensuring investor protection (Article 1(1) and Recital (10)). To achieve harmonization, the directive defined that a prospectus had to include a summary, a registration document with information regarding the issuer, and a securities note with information regarding the securities (Enriques and Gatti, 2008). To protect investors, prospectuses should "contain all information which, (...), is necessary to enable investors to make an informed assessment (...)" and this information should be "presented in an easily analyzable and comprehensible form" (Article 5(1)).

To implement the EU Prospectus Directive, the EU adopted Commission Regulation 809/2004 in 2004. It specified the design of prospectuses for different types of securities. An important specification was that prospectuses had to include a separate section on the risk factors related to the issuer and the security (Article 25(3)). Thus, we will separately examine the role of the risk factor section in our empirical analysis in Section V.5.

In 2010, according to the timetable laid out in the EU Prospectus Directive, the EU assessed the effects of the above-mentioned directive. To further reduce unnecessary obstacles for issuers and to strengthen investor protection, Directive 2010/73/EU of the European Parliament and of the Council ("Amending Directive") was adopted. A key goal was to increase the effectiveness of prospectuses by unifying their format (Fischer-Appelt, 2010). However, as the European Commission acknowledged in 2015, the 2003 directive and the reforms in 2010 were not successful. In particular, prospectuses were found to be too long and ineffective in protecting investors (European Commission, 2015).

To address these shortcomings, the EU adopted Regulation 2017/1129 ("EU Prospectus Regulation") in 2017 and thereby replaced the EU Prospectus Directive from 2003. The EU Prospectus Regulation came into force in 2019 and still applies. A central goal of this

[4]While Article 3(1) did not apply for securities offering to qualified investors (Article 3(2a)), as is typically the case for ABS, Article 3(3) was also binding for ABS. We observe that many ABS in our sample are indeed admitted for trading on a regulated market and thereby fall under the directive.

regulation, again, is to increase investor protection. Specifically, it requires in Recital (27) "that information should be sufficient and objective and should be written and presented in an easily analyzable, concise and comprehensible form. (...). A prospectus should not contain information which is not material or specific to the issuer and the securities concerned, as that could obscure the information relevant to the investment decision and thus undermine investor protection." In this way, the EU Prospectus Regulation specifically addresses the problem that standardized, complex, and long prospectuses are of little use to investors and can hinder an appropriate investment decision.

In addition to these regulations on the design of prospectuses for security issuances in general, there are also regulations specific to ABS. Most relevant to our study is Regulation 2017/2402 of the European Parliament and of the Council ("Securitization framework"), which was adopted in 2017 and has been in force since 2019. It established a set of new rules in the European securitization market that address the inherently complex and opaque nature of ABS and have the goal to "better differentiate simple, transparent and standardized products from complex, opaque and risky instruments" (Recital (3)). Accordingly, "all underlying documentation that is essential for the understanding of the transaction" should be made available to investors upon request to enable them to fully assess the risks associated with the investment (Article 7(1)). While the regulation contained many important aspects with respect to the transaction, these aspects should also be reflected in the information to be provided to investors. After the adoption of this regulation, the overall amount of information significantly increased, and a number of additional formal contracts had to be disclosed (Billio et al., 2023). This regulation is additionally tightened for those transactions that are intended to be "simple, transparent and standardized"-labeled (Article 22), for example, by further disclosure of statistical investigations on the loan portfolio prior to investors' pricing. We investigate the effectiveness of this regulation and the EU Prospectus Regulation in Section V.7.

II.2 Literature review

Theoretically, it is unclear how prospectuses affect investors in their security pricing. As derived in Section I, the issuer does not have the same motives as those of the investors and regulators when preparing the prospectus. Although prospectuses play an important role in the issuance process of securities and the sale of funds, they have been studied relatively scarcely. Therefore, an empirical investigation is required. Focusing on the literature on bond issuances and their prospectuses, Cai et al. (2022) find a relationship between textual risk disclosures in corporate debt issuances and the credit ratings of the bonds and the initial yield spread. With respect to ABS, Ghent et al. (2019) use prospectus characteristics of mortgage-backed securities (MBS) measuring the quantity of information such as the file size and the number of terms in the glossary as proxies for the complexity of the MBS. They find that complexity decreases security performance but is not priced properly by investors. Hibbeln et al. (2022) build on these findings and show that originators in the MBS market use the readability of prospectuses to obfuscate low security quality. Most importantly for our study, Zhang et al. (2023) analyze prospectuses and other written communication in residential mortgage-backed security (RMBS) offerings prior to the financial crisis. They find that such written communication contains important information regarding future security performance. Many investors, however, seem to neglect this information in their pricing. Naturally, this raises the question of why the textual information is not properly priced, and we address this question in our empirical analysis.

More broadly, previous literature examines how information provision in general, not just in prospectuses, affects the pricing of securities. Ertugrul et al. (2017) find that longer and more complex annual reports lead to higher firms' costs of debt. Bonsall and Miller (2017) confirm this finding regarding investors' pricing of debt securities and extend the analysis on ratings of the major rating agencies, which are also worse if the corporate disclosure is written especially less readable. Neilson et al. (2022) show that the provision of more granular quantitative information is beneficial for the accuracy of investors' ABS risk evaluation.

Beyond the literature on security issuances and corporate debt, there is a small but growing strand of literature that examines prospectuses in other contexts. deHaan et al. (2021) analyze mutual fund prospectuses and find that fund managers use long and complex prospectuses to obfuscate high fees. Abis et al. (2021) also focus on mutual funds and show that generic prospectuses confuse investors in their investment decisions and lead to greater fund size but also to higher fund flow volatility. Bourveau et al. (2022) explore prospectuses in the opaque and individualized market of ICO of crypto-tokens and find that lengthier and more technical prospectuses, positively interpreted as more information, lead to a higher amount of raised capital.

The studies examining prospectuses and security pricing are connected to a larger strand in the literature analyzing corporate and central bank disclosures. We focus on the three textual characteristics stressed by the European Securities and Markets Authority (2022) as important dimensions of the textual investor information and exploited in our study: the amount of boilerplate language, the linguistic complexity, and the length of the textual disclosure. With respect to boilerplate language and length, Cazier et al. (2021) find that firms with longer and more standardized risk factor disclosures experience favorable regulatory and judicial assessments, for example in the regulatory evaluation of the disclosure or in lawsuits against the firm. This important finding points toward firms being able to decrease regulatory-related efforts and legal risks by designing their textual disclosures accordingly. McClane (2019) addresses the idea that boilerplate language is used because it represents phrases that have already been tested in previous issuances. The study reveals that the use of boilerplate language is indeed associated with efficiencies in the preparation of the prospectus, but investors appear to penalize it. Kravet and Muslu (2013) show that a stronger increase in the length of textual risk disclosures in Form 10-K filings leads to a greater dispersion of investors' risk assessments, measured by higher stock return volatility, while it is less dispersed when the disclosure contains more boilerplate language.

In other contexts, Hansen et al. (2017) provide evidence that those members of the Federal Open Market Committee who are especially concerned about their reputation are incen-

tivized to use standardized language for precautionary purposes if the transparency of their deliberations raises, even though, overall, this effect is dominated by the disciplining effect of the rise in transparency. Kleymenova and Tomy (2022) emphasize that supervisors who are more concerned about negative consequences arising from the public's and market participants' reactions to their explanations tend to use more boilerplate language in accordance with the interpretation that standardized phrases are used to minimize potential risks from the interpretation of the communication. In line with the labeling of Hansen et al. (2017), boilerplate language in publicly disclosed firm reports is often interpreted as a proxy for textual disclosure quality, where more boilerplate language indicates more generic and standardized texts and thus lower textual disclosure quality (e.g., Lang and Stice-Lawrence, 2015; Abis et al., 2021).

With respect to linguistic complexity, Li (2008) provides evidence that companies with more linguistically complex annual reports hide detrimental information from investors. Lo et al. (2017) extend the analysis of Li (2008) by showing that companies that have to exceed their strong prior-year earnings write less readable annual reports. Bloomfield (2008) attributes the greater linguistic complexity in annual reports from firms that incur losses to the fact that these companies have to justify their poor results in greater detail. Miller (2010) argues that longer and more linguistically complex corporate disclosures are related to less trading activity, particularly by small investors. Lawrence (2013) shows that investors invest more in firms with less complex disclosures. The findings of Bushee et al. (2018) for quarterly earnings conference calls further indicate that a distinction must be made between whether complex language is used due to obfuscation or due to complicated content. Most recently, Bae et al. (2022) reveal that more complex Form 10-K filings increase uncertainty in the short run after the filings.

Overall, there is no unambiguous prediction in the existing literature as to how the characteristics boilerplate language and length of the prospectus are expected to affect the investors' risk assessment and pricing decisions, whereas linguistic complexity is mostly associated with negative consequences for investors and issuers of the publications.

III Data sources and sample selection

To comprehensively analyze the impact of the textual disclosure quality and quantity in ABS prospectuses on investors, we use a total of five different data sources. Our main data source is the EDW. It has collected, validated, and disseminated securitization data at the deal, security, and loan level since 2012. Our data covers all transactions reported to the EDW and issued between 2002 and 2020.[5] The ABS in our sample include loans from Austria, Belgium, France, Finland, Germany, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, and the United Kingdom. Overall, our data set represents almost all European countries active in securitization (Association for Financial Markets in Europe, 2014).

We rely on this data source for issuance prospectuses as well as for deal-, security-, and loan-level information. We downloaded the prospectuses from the database of EDW and then manually supplemented data gaps with publicly available prospectuses. For an additional analysis, we manually extracted the sections of the prospectuses where the risk factors for investors are described. In total, 1,188 prospectuses across six asset classes were available. The asset classes include automobile loans, consumer loans, credit card loans, leasing contracts, residential mortgage loans, and loans to small- and medium-sized enterprises (SME).[6] Details of the ABS transaction distribution by asset class and country in our sample are provided in Figure A.1 in the Internet Appendix.

The second important data source supplementing our data set is S&P Global, which provides performance, rating, and pricing data on the ABS. These data comprise the initial yield spread demanded in the issuance process, traded spreads in the secondary market, as well as security ratings and performance information over the security lifetime. Moreover,

[5]Due to the special character of commercial mortgage-backed securities (CMBS) transactions, the respective loan portfolios consist of a small number of very large, mostly publicly known loan contracts. Consequently, we do not include this asset class in our analysis. However, the absolute number of CMBS transactions for which loan-level data are reported to EDW is only five.

[6]If we re-run our analyses of Section V separately for each asset class, our results remain predominantly in line with our main results.

we add interbank lending rates and sovereign bond spreads for the calculation of yield spreads from three data sources, namely the ECB, FRED, and Refinitiv Datastream.

As we are interested in the ABS pricing at the time of issuance, we only use data available for the year of the respective ABS issuance and, consequently, only one observation per security. Our entire sample selection and preparation process is summarized in Table A.1 in the Internet Appendix. We start with 6,691 security observations, belonging to 1,406 ABS deals. In the first adjustment step, we exclude deal observations for which no prospectus information is available. In the second step, to ensure consistent calculation of our measures of textual disclosure quality and quantity, which are based on the written content of the prospectuses, we exclude all prospectuses that are not written in the English language. In the third step, we exclude deal observations where the issuer is not clearly identifiable. After these adjustment steps, 4,301 observations of securities remain, which represent 1,014 different ABS portfolios, including more than 39.4 million single loan observations.

The 1,014 ABS deals represent our final deal-level sample, which we exploit in our analysis of recent regulations (Section V.7). For our analysis on the security level, we need additional information, and thus, exclude all observations for which ratings, coupon prices, and interest shortfall amounts are missing. Finally, and in line with the previous literature (e.g., Ertan et al., 2017; Hibbeln and Osterkamp, 2020; Klein et al., 2021), we also exclude implausible observations in the security level data. Dropping these observations and also those for which the further control variables are not available results in a final security-level sample of 2,469 securities from 866 ABS deals.

IV Variables and descriptive statistics

IV.1 Measuring textual disclosure quality and quantity

In this section, we define the variables used in our empirical analysis. We first introduce the variables measuring the textual disclosure quality and quantity in the prospectuses

and present descriptive statistics on these. Thereby, we provide important insights into the textual design of the issuance prospectuses in our data set. In the second part of this section, we further describe variables related to ABS pricing and performance, as well as our control variables. In addition, brief definitions of all variables can be found in Table A.2 in the Internet Appendix. Summary statistics based on the deal-level and security-level samples can be found in Table 1. Pairwise correlation coefficients are presented in Table A.3 in the Internet Appendix. In line with the studies of Easley and O'Hara (2004) and Kleymenova and Tomy (2022), we focus on three key variables of the prospectuses' textual disclosure quality and quantity characteristics: *Boilerplate language*, *Linguistic complexity*, and *Prospectus length*.

[Table 1 about here.]

As our first measure, we quantify the share of *Boilerplate language* in ABS prospectuses following the idea initially proposed by Lang and Stice-Lawrence (2015) and subsequently used by Dyer et al. (2017), Cazier et al. (2021), Kleymenova and Tomy (2022), and others. Thereby, we measure the extent to which the prospectuses consist of standardized phrases, potentially indicating generic and, therefore, uninformative rather than individual descriptions. We define boilerplate phrases as four-word phrases (tetragrams) that are very common in the prospectuses within the same asset class. For this step, we follow Kleymenova and Tomy (2022) and exclude common words and terms like "and", "the", or "Basel committee". We then define all tetragrams appearing in at least 33.33% of the prospectuses within the same asset class as boilerplate phrases.[7] Table A.4 in the Internet Appendix presents some of the boilerplate phrases we identify and shows that these phrases are indeed generic and are not combined with security-specific numeric information. To finally quantify *Boilerplate language*, we then measure the number of words appearing in

[7]We do not include tetragrams appearing in 100% of the prospectuses of ABS deals within the same asset class. This is based on the idea in the literature that such tetragrams are often innocuous or based on regulation (e.g., Lang and Stice-Lawrence, 2015).

sentences with boilerplate phrases in a prospectus relative to the total number of words in the prospectus.[8]

Second, we quantify *Linguistic complexity* as a measure of an easily understandable and comprehensible form of information presentation. Thus, we calculate Gunning's Fog index of the prospectuses, a well-established measure of linguistic complexity commonly used in this context (e.g., Li, 2008; Miller, 2010; Lang and Stice-Lawrence, 2015; Lo et al., 2017; Bushee et al., 2018; Kleymenova and Tomy, 2022). Gunning's Fog index is defined as 0.4 x (average sentence length + 100 x (number of complex words / number of words)).[9] The index has a convenient interpretation because it represents the number of years of learning in the US education system, required for an average reader to comprehend the content of a text after reading it for the first time (Li, 2008). Therefore, a higher value indicates lower readability, i.e., higher linguistic complexity of the text.

Loughran and McDonald (2014) have initiated a discussion about how to appropriately measure readability. They argue that, for standardized annual reports, instead of traditional linguistic measures like Gunning's Fog index, the size of the document is more likely to accurately represent the text's readability. In contrast to Loughran and McDonald (2014), Bonsall et al. (2017) point out that when using file size as a measure of readability in financial disclosures, the results may strongly be driven by figures, images, and other characteristics of the document which are not related to the linguistic style of the text. Using "Gunning-Fog and Flesch-Kincaid models to judge the level of compliance with the plain English rules" is also spurred by the Securities and Exchange Commission (SEC) and its former chairperson, Christopher Cox (Cox, 2007). We consider the arguments by Bonsall et al. (2017) to be particularly important in the context of prospectuses, which typically contain several figures and tables, and follow the SEC proposition in this paper.

[8]To increase the comparability of the regression coefficients in Section V, we scale this variable so that it ranges from 0 to 1, reflecting a share of 0 to 100 percent. All other variables expressing shares range from 0 to 100.

[9]In Gunning's Fog index, complex words are defined as words that consist of three or more syllables. We further exclude sentences with more than 100 words to avoid potential parsing problems affecting our measure. However, our results are robust to omitting this restriction.

Third, we consider *Prospectus length* as our third measure of the prospectuses' textual disclosure quality and quantity, quantified as the natural logarithm of the total number of words in a prospectus. We measure the prospectus length to address the argument that prospectuses could be too long and, in this way, deter investors from reading them or finding the most relevant information. The length of corporate disclosures is frequently examined in the literature and typically measured as in our study (e.g., Li, 2008; Lang and Stice-Lawrence, 2015; Cazier et al., 2021).

To ensure that our results are not overly dependent on the specifications of the three measures, we also use alternative definitions in our empirical analysis. For boilerplate language, we use a more conservative specification based on six-word phrases that appear in at least 50% of prospectuses from ABS deals of the same asset class. For identifying linguistic complexity, we use the also commonly applied Flesch-Kincaid index, and for measuring prospectus length, we use the logarithmized number of the prospectus pages.

To get a better understanding of the textual characteristics of the ABS prospectuses in our sample, we first examine the distributions of the three textual disclosure measures. Figure 1 presents histograms for all three measures, based on our deal-level sample.[10] Interestingly, we see that *Boilerplate language* exhibits a wide range from below 10% up to more than 70%. This indicates that some prospectuses include much more standardized phrasing than others. As displayed in Table 1, the sample mean is 43% (and 39 % at the security level). When turning to *Linguistic complexity*, we observe values ranging from 20 up to 26 with a sample mean of 23.95 (and 23.73 at the security level). Given the interpretation of the index as the years of formal education necessary to understand the text after the first reading, these values are remarkably high and correspond approximately to a completed master's degree or PhD. Furthermore, a difference of six years is quite substantial. The variation can also be seen in Figure A.5 in the Internet Appendix. It shows that the description of a similar risk, specifically the risk of early repayments, can be written with a Gunning's Fog index of 19.52 or a Gunning's Fog index of 28.65. With respect to *Prospectus length*, we

[10]We also examine the distributions of the three measures for each asset class individually. The results are presented in Figure A.2, Figure A.3, and Figure A.4 in the Internet Appendix.

observe values ranging from close to 39,000 words to more than 170,000 words. The mean of *Prospectus length* is close to 92,000 words or 11.43 in logarithmic terms (and 11.42 at the security level). Taken together, these insights provide first evidence that prospectuses are, on average, very standardized and generic, hard to comprehend, and very long but vary substantially in these characteristics.

[Figure 1 about here.]

Next, we examine the development of the three ABS prospectus design measures over time. Figure 2 plots the mean values as well as intervals of one standard deviation above and below the mean of the measures per year. With respect to *Boilerplate language*, we observe an increase from 2004 to 2011 and a subsequent consolidation on a level close to 50%. Furthermore, the dispersion around the mean is somewhat smaller in the 2010s than in the 2000s. With respect to *Linguistic complexity*, prospectuses have become increasingly complex over most of the sample period. Turning to *Prospectus length*, we observe that prospectuses have become increasingly long, especially after 2010. However, this trend appears to be reversing recently. Interestingly, we observe that the mean values of all three measures follow a similar pattern and fall in 2010 after the financial crisis but then quickly return to and even exceed pre-crisis levels. Therefore, overall, the issuance prospectuses have become substantially more standardized, more complex, and longer over the sample period. This provides additional motivation for our study.

[Figure 2 about here.]

IV.2 Pricing, performance, and control variables

ABS pricing and performance:

To evaluate investors' security pricing and their expectations regarding the security default risk and its development during the term, we use the *Yield spread* at security issuance.

For well-informed investors, *Yield spread* should approximately reflect the actual security risk. The very granular pricing data in our study obtained from S&P Global and EDW allow us to use the initially demanded yield spreads per security instead of relying on averaged yield spreads per deal. In general, a distinction is to be made between fixed and floating coupon payments. Following Mählmann (2012), we incorporate for both, variable and fixed coupons, the coupon amount at security issuance. While *Yield spread* is defined as security coupon payments above the reference interest rate if the coupon is floating, we follow He et al. (2016) and deduct the maturity-adjusted risk-free rate from the coupon at issuance if the coupon is fixed.[11] As risk-free rate, we use the ECB yield spread index of all sovereign bonds which are "AAA"-rated in the Euro area. The development of the *Yield spread* over time is presented in Figure 3. It shows that the average *Yield spread* is rather constant and slightly below the value of 1, in line with the mean value of 0.91 as presented in Table 1. Exceptions of higher yield spreads are in the years 2009 to 2012, the period after the financial crisis and during the Euro area sovereign debt crisis, as well as in 2019 to 2020, reflecting higher default and interest rate risk expectations of the investors.

[Figure 3 about here.]

In the second part of our analysis, we examine the impact of textual disclosure quality and quantity on investors' ability to adequately assess securities' risk. Following the commonly used approach in the literature (e.g., Becker and Milbourn, 2011; Mählmann, 2012; He et al., 2016; Bonsall and Miller, 2017; Neilson et al., 2022), we estimate how textual disclosure quality and quantity in ABS prospectuses affect the predictive ability of initial yield spreads for future security performance. To quantify the ex-post performance of a security, we use the *Interest shortfall*, which is defined as the following fraction: The numerator is the cumulative maximum interest shortfall amount during the securities maturity, while the denominator is defined as the principal balance of the security at the time of issuance to control for the security size. Herein, the securities' interest shortfall is the difference

[11] In order to comprehensively show that our results are not driven by the underlying coupon type, we include coupon type fixed effects and robustness checks using only the significantly larger part of our data set, i.e., the securities with floating coupons. See Sections V.1 and V.2 and Tables A.7 and A.9 in the Internet Appendix for further details. Overall, our results remain qualitatively unchanged.

between the contractually agreed securities' interest payment and the actual interest payment. If this value is greater than zero, investors have not received the full interest payment at some point during the securities' lifetime.[12]

Security- and deal-level control variables:

In addition to our key variables, we include a broad set of control variables, all measured at security issuance. To take the credit risk of the underlying loans in the securitized portfolio into account, we control for the *Interest rate* as the average interest rate in the loan portfolio. We incorporate the structural complexity variables *Number of securities* in an ABS deal and *Rating disagreement*, defined as an indicator variable is 1 if the three most important credit rating agencies (CRAs), Fitch, Moody's, and S&P, did not agree on at least one security rating in the respective ABS deal and 0 otherwise. [13] The *Security width* indicates the share of a security as a percentage of the total volume of a deal. We also control for the absolute *Security size*, measured as the natural logarithm of the principal balance of the ABS, because larger securities in ABS deals can offer a higher degree of risk diversification and increased liquidity, accompanied by lower yield expectations and, therefore, lower shortfall rates (Peña-Cerezo et al., 2019). We calculate *Security term* as the natural logarithm of the period between the security issuance and the planned security maturity date expressed in years as longer maturities may characterize safer securities (Helwege and Turner, 1999). Following Mählmann (2012) and Ghent et al. (2019), we include the amount of *Excess interest* and *Subordination* as a form of credit enhancements as further controls. *Excess interest*, also called excess spread, refers to the difference between received payments by the security's issuer and the interest paid to the investors. *Subordination* of a security is defined as the percentage of ABS deal volume that is subordinated to the respective security, and thus, it suffers losses when the corresponding percentage is exceeded (Mählmann, 2012).

[12]It is possible that investors did not receive the contractually agreed upon payment at the contractually agreed upon time but received a back payment later so that they eventually did not have to bear losses. In this case, our measure reflects the maximum cumulative delayed interest shortfall.

[13]We also alternate the definition of *Rating disagreement* and re-estimate our analyses using a variable definition relating only to the individual security. In this case, the value of the indicator variable *Rating disagreement* of a security does not depend on the rating of the other securities in the same deal. Our results remain qualitatively unaffected.

V Effects of textual disclosure quality and quantity on investors

V.1 Textual disclosure quality and quantity and the initial yield spread

In the first step of our empirical analysis, we investigate whether and how the three previously introduced textual disclosure measures in the prospectuses, *Boilerplate language*, *Linguistic complexity*, and *Prospectus length*, affect the level of the risk premium investors demand for the securities in the issuance process.

Empirical strategy:

We estimate a linear regression model and incorporate the initial *Yield spread* as the dependent variable, and *Boilerplate language*, *Linguistic complexity*, and *Prospectus length* as our independent variables of main interest, which we include both individually and jointly in the estimation. Thus, we calculate the following ordinary least squares (OLS) regression equation:

$$\begin{aligned} Yield\ spread_{i,s} = \beta_0 + \beta_1\ x\ Text.\ discl.\ quality\ and\ quantity_i \\ + \beta_2\ x\ Controls_{i,s} + \alpha_i + \gamma_i + \zeta_i + \kappa_s + \rho_s + \epsilon_{i,s} \end{aligned} \tag{V.1}$$

In the notation used in Equation V.1, i indexes ABS deals and s indexes securities. $\epsilon_{i,s}$ is an error term. To isolate the effect of textual disclosure quality and quantity on the initial yield spread, we control for several security-specific characteristics that are likely linked to ABS pricing (see Section IV.2 for further details on the control variables). Additionally, we incorporate five different types of fixed effects (FE). First, we include deal origination year FE (α) to control for unobserved dynamics connected with the security origination year. Second, we use asset class FE (γ) to control for unobserved variation in security performance across the six incorporated asset classes and, third, country FE (ζ) to avoid our results being driven by unobserved heterogeneity corresponding to the country of the underlying collateral, which is usually identical with the country of the underlying loans. Fourth, we include rating FE (κ) to control for varying issuance ratings of ABS reflecting

external expectations regarding their risk and future performance. Fifth, we add coupon type FE (ρ) to control for unobserved heterogeneity with respect to floating or fixed coupon types of the ABS. Standard errors are clustered at the ABS deal level because we observe different securities for each deal and, therefore, account for correlation within an ABS deal.[14]

Results:

The results are presented in Table 2. The coefficient of *Boilerplate language* is negative and highly statistically significant, clearly indicating that investors perceive standardized language, i.e., phrases that are very commonly used across ABS prospectuses, as a positive signal and demand lower yield spreads at issuance. The coefficient of 0.830 further suggests that an increase of *Boilerplate language* by one standard deviation relates to a decrease in the initial yield spread of about 17 bp (0.166 = 0.830 x 0.20), which represents about 18% of the mean initial yield spread in our sample. Therefore, the relationship is also economically significant. In combination with recent findings by Cazier et al. (2021) that regulators and courts react positively to boilerplate language as well, our results suggest that issuers benefit from the use of boilerplate language even though it reflects a less individual disclosure and thus potentially hinders an accurate risk assessment for a specific security—an issue that will be examined in Section V.2.

[Table 2 about here.]

Turning to *Prospectus length*, we find a positive and statistically significant relation. This indicates that investors penalize ABS with longer textual descriptions by demanding a higher initial yield spread. With respect to the effect size, an increase in the logarithmized number of words in the ABS prospectus by one standard deviation relates to an increase in the initial yield spread of 0.115 (0.311 x 0.37) or 12 bp, which makes up around 12.6% of the mean initial yield spread and is again economically meaningful. Consequently, higher

[14] Our results (of this analysis and also of the subsequent ones) remain qualitatively unchanged if we alter the clustering dimension(s) and, for example, re-estimate the regression equation using double-clustering at the levels of the deals and the deal origination years.

textual disclosure quantity does not seem to pay off for issuers with respect to the security price. Together with the above-mentioned findings, Cazier et al. (2021) also show that not only a higher share of *Boilerplate language* but also a higher textual disclosure quantity decreases legal risks, resulting in a trade-off for issuers between the investors' demand for the *Yield spread* and the legal perspective. With respect to *Linguistic complexity*, we observe a negative yet insignificant coefficient, indicating that the relationship between the readability of the prospectus and the initial yield spread demanded by investors is not as straightforward as compared to *Boilerplate language*. Finally, when all three measures of ABS prospectus design are included together in the regression, our results remain qualitatively unchanged.

Furthermore, we analyze which bank characteristics induce the use of *Boilerplate language*, *Linguistic complexity*, and *Prospectus length*. For this analysis, we regress our textual characteristics on five main bank characteristics: bank size, liquidity, ROE, Tier-1 capital ratio, and the NPL ratio.[15] As shown in Table A.5 in the Internet Appendix, larger banks and those with a better liquidity position write, on average, longer prospectuses. In contrast, those with a low regulatory equity but a high NPL ratio write especially complex prospectuses containing a large portion of standardized language. These relations underpin that the investors' reactions to the textual measures are not driven by presumed correlations of the textual measures with the risk of the average bank loan portfolio, as shown by the coefficients of the NPL ratio. This finding is also affirmed when including the bank characteristics as controls in our main regression.

To ensure that the results are not driven by the specification of the textual disclosure measures, we re-run the analysis using different specifications of the measures as discussed in Section IV.1. These are a boilerplate measure based on six-word phrases that appear in at least 50% of the prospectuses, the Flesch-Kincaid index as a measure of linguistic complexity, and the logarithm of the number of pages in the ABS prospectus as a measure of prospectus length. The results are presented in Table A.6 in the Internet Appendix

[15]The regression is conducted at the deal level and resembles Equation V.3 of Section V.7.

and remain qualitatively unchanged, documenting the robustness of our findings. Furthermore, we estimate the regression equation only including ABS with floating coupons, and again, our results remain qualitatively unchanged (presented in Table A.7 in the Internet Appendix).

Mechanisms and insights:

In this paragraph, we provide further insights into which ABS are especially affected by the impact of the textual disclosure quality and quantity on the initially demanded *Yield spread* and which mechanisms underlie our findings above. For this, we re-estimate Equation V.1 and expand it by interacting our textual disclosure measures with the variables *Rating*, *Deal count*, and *Visualization* as measures of the investors' presumed security risk, the information uncertainty on the respective deal or bank, and additional visual presentations of the descriptions, terms, and conditions in the prospectus. Table 3 presents the results on these determinants driving the investors' perception of the security risk and the respective initial security pricing.

[Table 3 about here.]

First, we take advantage of a particularity of the ABS deal structure in our sample. The default risk and the corresponding credit ratings of the issued securities assigned by the credit rating agencies exhibit a peculiarly large dispersion due to the securities' individual level of subordination within the cascade system (Furfine, 2014). We define *Rating* as a linear expression of the average rating assigned by one (or more) of the credit rating agencies (S&P, Moody, and Fitch).[16] The results of the interaction term of *Rating* with our textual disclosure measures are presented in columns (1), (4), and (7) of Table 3 and provide three positive and significant coefficients of the interaction, while *Boilerplate*

[16]This variable is defined in line with Cai et al. (2022) as follows: "AAA" is transformed into the numerical value 1, "AA+" into 2 up to "C" into 19, and finally "D" into 20. We choose the linear transformation of the ratings into numbers in order to receive regression coefficients, which are easier to interpret. We re-estimate all regressions using a non-linear transformation of the ratings into the respective probability of default of the rating categories and our results remain qualitatively unchanged.

language and *Linguistic complexity* have a significant negative coefficient. Due to our rating FE, the individual coefficient of *Rating* is omitted.

The results show that the negative impact of *Boilerplate language* on *Yield spread* arises from securities with high ratings, that is, low default risk. The positive coefficient of *Boilerplate language x Rating*, however, demonstrates that this effect decreases with increasing default risk. Calculating the net effect of *Boilerplate language* and *Boilerplate language x Rating* shows that the impact of *Boilerplate language* on the demanded *Yield spread* is negative up to the rating "BB-", approximately zero for the ratings "BB-" and "B+" and reverses with lower ratings. The result emphasizes that *Boilerplate language* is perceived as a signal of lower but even stronger as a signal of a standardized security's risk profile without extraordinary and individual risks of an economic breakdown as the investors of senior tranches, often "AAA" rated, largely depend on the overall macroeconomic crash risk (like the Global financial crisis or the Covid pandemic) instead of the individual transaction characteristics. On the opposite, investors of ABS with lower seniority and lower ratings are more prone to suffer losses and thus expect these losses not only in economically extraordinary crises and are thus more incentivized to focus on the individual description and unambiguous information of the risks and the transaction structure (Neilson et al., 2022).

This mechanism is also shown when focusing on *Linguistic complexity*. In line with the literature on linguistic complexity in sales prospectuses (see II.2), a higher level of *Linguistic complexity* decreases the *Yield spread* for securities with high ratings, but this effect decreases with a higher level of assumed security default risk and also reverses for lower ratings.

Finally, we find a significantly positive coefficient on the demanded risk premium for the interaction term *Prospectus length x Rating*, whereas the coefficient of *Prospectus length* is no longer statistically significant. This indicates that the investors respond particularly sensitively to the length of the descriptions in the issuance prospectuses when the security default risk is substantial and they expect to suffer future losses. As the variable *Rating*

exhibits a minimum value of one, the net effect of the *Prospectus length* and its interaction with *Rating* on *Yield spread* is positive for all rating categories, consistent with our previous finding on the impact on *Prospectus length*. Unreported, we find the same results examining the security eligibility in the ECB repo framework instead of *Rating*. This shows that the ECB eligibility itself does not affect our results beyond the ECB's rating requirements in its collateral framework.

Second, *Deal count* provides the number of securitization deals a bank issued in our data set. As investors may learn from previous deals of the same bank about the bank's behavior and loan portfolio characteristics, such as credit origination or monitoring practices, the level of asymmetric information and uncertainty on the future repayments differs depending on *Deal count*, with higher numbers showing lower level of uncertainty. Consequently, investors depend stronger on information from the issuance prospectuses of those securities, at which the respective bank issued fewer transactions.

Column (2) of Table 3 shows that a higher level of *Boilerplate language* leads to a lower demanded *Yield spread*, but this effect largely diminishes with an increasing *Deal count*, while the net impact of *Boilerplate language* and *Deal count* on *Yield spread* remains still negative even for the most active bank in our sample (maximum *Deal count* is 39). Thus, we show that *Boilerplate language* is particularly impactful and particularly changes investors' perceptions, whether the securities are exposed to high or exceptional risks, when investors' uncertainty is high. This finding is in line with the theoretical arguments by Hansen et al. (2017) and Cazier et al. (2021) on the precautionary purposes as incentives for standardized language.

Regarding *Linguistic complexity* and *Prospectus length*, positive interaction terms of *Linguistic complexity* and *Prospectus length* with *Deal count* in columns (5) and (8) show an increasing positive impact of lengthier (more complex) prospectuses on *Yield spread* if the bank frequently securitizes. Calculating the net effect of *Prospectus length* (*Linguistic complexity*), *Deal count*, and their interaction, we find an almost unified negative impact of these coefficients if *Prospectus length* is low, regardless of the level of *Deal count* as well

as for a higher *Prospectus length* in combination with a low *Deal count*. However, we find only a marginally negative or even positive net effect if the bank frequently securitizes and uses a lengthy description in the prospectuses. This indicates that lengthier prospectuses are seen as a warning signal by investors if the bank is regularly present on the market, information asymmetries are low, and there should, therefore, be less need for extensive information.

Third, we measure the extent of *Visualization* in the prospectus by the logarithmized file size (in mb).[17] While a higher level of *Visualization* allows easier and faster prospect processing by investors, it also shortens information and makes it harder to describe complex and specific contents (e.g., Lurie and Mason, 2007). The results, presented in columns (3), (6), and (9) of Table 3, show that the impact of our textual disclosure measures on the demanded *Yield spread* is strongly affected by the level of *Visualization* in the prospectus. Calculating the net effects of the coefficients of *Boilerplate language*, *Visualization*, and *Boilerplate language x Visualization* shows that, if the level of *Visualization* is high (mean plus one standard deviation, 0.4+0.6=1), an increase of *Boilerplate language* by one standard deviation lowers the *Yield spread* by about 21 bp, while the increase in *Boilerplate language* lowers the *Yield spread* by only 6 bp when the level of *Visualization* is low (0.4-0.6=-0.2). As this relation similarly holds inversely when varying *Visualization* given a high (low) level of *Boilerplate language*, this result reveals a complementary relation between *Boilerplate language* and *Visualization*. Additionally, we find that the increasing impact of *Prospectus length* on *Yield spread* is weakened by *Visualization*, as shown by the negative interaction term.

[17] We use this indicator in line with the argumentation of Bonsall et al. (2017) that emphasize that the file size is strongly determined by figures, images, and other representations within the document, which have a much higher demand for data storage than plain text. This is especially prevalent in the case of the very individualized prospectuses in our sample, so the extent of graphical illustrations differs substantially. For the discussion on the file size of Loughran and McDonald (2014), see II.2.

V.2 Textual disclosure quality and quantity and the accuracy of the initial yield spread

In the second step of our empirical analysis, we investigate the effect of textual disclosure quality and quantity on the accuracy of investors' risk assessment. Therefore, we analyze how the textual disclosure measures affect the ability of the yield spread demanded by investors at issuance to predict future ABS performance, beyond the assigned credit ratings.

Empirical strategy:

Our empirical strategy follows a number of recent studies investigating the effect of asset-level transparency, structural deal complexity, and other factors affecting the predictive ability of yield spreads (e.g., Mählmann, 2012; He et al., 2016; Neilson et al., 2022). In this setting, we incorporate the ex-post performance of an ABS (*Interest shortfall*) as the dependent variable. Our independent variables of main interest are the *Yield spread* and its interaction with our measures of textual disclosure quality and quantity. The idea behind this setting is that investors assess the risk of an ABS at security issuance and price this risk in their demanded initial yield spread for the ABS. If the investors' risk assessment is accurate, the initial yield spread—as a proxy for the risk assessed by the investors—should be a good predictor of future ABS performance because a higher security risk should translate into a higher yield spread and a higher interest shortfall. This mechanism is displayed in Figure 4 in the Appendix. Based on the interaction term, the setting then allows us to examine how the textual disclosure quality and quantity affect the predictive ability of initial yield spread. A significant coefficient of the interaction term (*Yield spread x Text. discl. quality and quantity*) indicates that the accuracy of the initially demanded yield spread and thus the investors' risk assessment (provided by (R) in the figure) is influenced by the textual disclosure in the issuance prospectuses (D). If the coefficient is negative, a higher initially demanded yield spread does not predict a higher interest shortfall appropriately (i.e., a lower ABS performance) and, thus, the initial investors' security pricing is distorted.

Besides these independent variables of main interest, we also include the measures of textual disclosure quality and quantity directly and again control for several security-specific characteristics and the five different FE as described above. Thus, we estimate the following linear regression model:

$$\begin{aligned} Interest\ shortfall_{i,s} = \beta_0 &+ \beta_1\ x\ Yield\ spread_{i,s} \\ &+ \beta_2\ x\ Text.\ discl.\ quality\ and\ quantity_i \text{ x } Yield\ spread_{i,s} \\ &+ \beta_3\ x\ Text.\ discl.\ quality\ and\ quantity_i \\ &+ \beta_4\ x\ Controls_{i,s} + \alpha_i + \gamma_i + \zeta_i + \kappa_s + \rho_s + \epsilon_{i,s} \end{aligned} \tag{V.2}$$

To estimate the regression coefficients, we again use an OLS estimator. As before, standard errors are clustered at the ABS deal level.

Results:

We present our results in Table 4. The coefficients of *Yield spread* are significantly positive across all specifications, indicating that investors seem to be capable of assessing and pricing the actual risks in ABS beyond the evaluation of the rating agencies. Most interestingly, the coefficients of the interaction terms are significantly negative across all three textual disclosure measures. This indicates that the predictive ability of initial yield spreads on future ABS performance is lower for deals where the ABS prospectuses are more standardized, complex, and longer. When examining the three interactions jointly, we find that the coefficient for the interaction between *Boilerplate language* and *Yield spread* remains statistically significant.

[Table 4 about here.]

To provide more economic intuition for these results and to outline their consequences for investors as well as the aggregate market, Figure 5 presents the predictions for the *Interest shortfall* based on the *Yield spread* for different levels of the three textual disclosure quality and quantity measures. The presented levels are the mean and its variation by one

standard deviation.[18] The estimates are calculated for an "A" rated, 2013-issued security with a floating interest rate and backed by a Spanish RMBS portfolio.[19] In the upper part of the figure, the *Yield spread* demanded by investors is positively correlated with the ex-post *Interest shortfall* only in case of a low or medium level of *Boilerplate language* but negatively in case of a high level of *Boilerplate language*. Thus, when prospectuses are written in a more standardized way, higher initially demanded yields do not predict higher shortfalls but, in fact, lower shortfalls. This means that investors demand too low yields for securities that perform bad ex-post and too high yields for securities that perform well ex post. Although this pricing behavior is not always to the disadvantage of investors, from a more general perspective, it is very detrimental to the effective functioning of securities markets. Moreover, these results are qualitatively similar when focusing on the linguistic complexity and the length of prospectuses. Therefore, the results suggest that particularly complex, standardized, and lengthy prospectuses are obstructive to the accuracy of investors' security pricing. This is in line with the intuition that such prospectus characteristics deter investors from understanding or possibly even reading information that is important for the accurate pricing of risk, for example, because the information is not specific enough or written in a too complex or long way.

[Figure 5 about here.]

Our results fit well with arguments from the literature on search costs (e.g., Ellison and Ellison, 2009; Ellison and Wolitzky, 2012). The idea in these arguments is that for more complex products, investors need to spend more resources when trying to understand the product. In this context, our results raise the question of whether investors are not able to predict the security default risk more accurately in advance due to a lower textual disclosure quality and higher quantity or whether it is not economically worthwhile for them to do

[18] More general, Figure A.6 in the Internet Appendix provides marginal effects of the *Yield spread* on the *Interest shortfall* for a broader and continuous range of values for the three textual disclosure quality and quantity measures.

[19] We choose a representative combination of the characteristics, for which we control by means of fixed effects, for this illustration. The graph and the respective description remain qualitatively unchanged for other combinations of our fixed effects variables.

so, since the search costs exceed the assumed benefits of the prediction improvements. The security volumes in the European ABS market are typically high, and there are only institutional investors in this market (in accordance with ongoing regulation, it is prohibited to sell ABS to retail customers in Europe). In line, anecdotal evidence confirms that typical investment volumes make even costly risk assessments worthwhile. However, both potential explanations point to an important loss of economic efficiency in capital markets.

To test the robustness of our results, we re-run the regressions again using our previously introduced alternative definitions of disclosure quality and quantity. The results are presented in Table A.8 in the Internet Appendix and remain predominantly unchanged. We further run a robustness check, including only ABS with floating coupons in our analysis. The results are presented in Table A.9 in the Internet Appendix and again remain qualitatively similar.

To further investigate the nexus between the accuracy of the initial yield spreads demanded by investors and the disclosure quality and quantity in ABS prospectuses, we also analyze whether the initial difficulties with adequately assessing risk lead to increased secondary market volatility after issuance as investors obtain ABS performance information from realized repayments over time and potentially adjust their pricing. For this analysis, we re-estimate the regression model from Equation V.1 but replace the yield spread as the dependent variable with the volatility of the secondary market spread (*Spread volatility*). *Spread volatility* is defined as the average daily change in the spread traded on the secondary market. The results of this analysis are presented in Table 5. Interestingly, we find that lengthier and linguistically more complex ABS prospectuses relate to a significantly higher secondary market spread volatility. The coefficient for *Boilerplate language* is positive as well but on the borderline of statistical significance. Hence, lower disclosure quality and higher disclosure quantity in ABS prospectuses are related to a higher volatility of secondary market spreads of the ABS. We interpret our findings in the sense that investors obtain new information on the risk and performance of the ABS over time and need to

adjust their initial pricing more extensively, thus adding to the volatility of secondary market spreads by including this information.

[Table 5 about here.]

Overall, the results of the second step of our empirical analysis suggest that more standardized, complex, and longer ABS prospectuses negatively relate to the accuracy of investors' risk assessment. This interpretation is well in line with the finding that for these prospectuses, the price adjustments after the ABS issuance are substantially higher.

V.3 Textual disclosure quality and quantity and the ABS performance

A possible concern for our results might be that our measures of textual disclosure quality and quantity of an ABS prospectus contain information on the security risk and performance that influence the initial security pricing and correlate with subsequent shortfalls, but they are not captured by our control variables. In this case, the impact of the disclosure characteristics on investors' initial security pricing, as provided in Section V.1, does not rely on the actual prospectus characteristics but on the included information on risk and performance in these measures. Additionally, the inclusion of textual disclosure quality and quantity in investors' initial security pricing would be reasonable and economically beneficial. However, this assumption does not correspond to our results of Section V.2 since additional, reliable information should lead to better pricing precision.

We follow the procedure of Zhang et al. (2023) and address the concerns by orthogonalizing *Boilerplate language*, *Linguistic complexity*, and *Prospectus length* to the subsequent security performance *Interest shortfall*. In the first step of this two-stage estimation procedure, we regress *Interest shortfall* as the independent variable on our textual disclosure quality and quantity measures as the dependent variable, respectively. We then use the residuals from each of these regressions as adjusted independent variables of main interest in our regressions and revisit the central findings of our previous analyses. Thereby, we measure

the textual disclosure quality and quantity in ABS prospectuses beyond the underlying ABS deal risk and performance. Our results presented in this analysis also hold for various variable definitions of initial ABS portfolio risk and subsequent security performance used in the orthogonalization procedure. Importantly, this procedure does not fall under the critique of Chen et al. (2018) since we use the newly generated, orthogonalized variables as the independent variable rather than the dependent variable in our second-stage estimation.

The results with respect to the effect of textual disclosure quality and quantity on the initial yield spread, as previously discussed in Section V.1, are presented in Table 6. The results remain qualitatively unchanged: The coefficient of *Boilerplate language* is negative and statistically significant, the coefficient of *Prospectus length* is positive and statistically significant, and the coefficient of *Linguistic complexity* remains statistically insignificant. We also revisit the results on the effect of textual disclosure quality and quantity on the predictive ability of yield spreads, as previously discussed in Section V.2. The results are presented in Table 9 and again remain qualitatively unchanged. Consequently, this analysis shows that the impact of the linguistic prospectuses' characteristics is not driven by unobserved correlations of the textual disclosure measures and portfolio and security risk and performance.

[Tables 6 and 7 about here.]

V.4 Textual disclosure quality and quantity and the complexity of ABS deals

Next, we address the idea that the textual characteristics of an ABS prospectus simply reflect the characteristics of the underlying ABS deal. This idea is inspired by Bushee et al. (2018), who disentangle linguistic complexity due to obfuscation from linguistic complexity due to complex information. For example, a longer and more complex prospectus might be the result of a more complex ABS deal structure. To address this concern, we re-estimate

the two-stage regression approach introduced in Section V.3 using ABS deal characteristics that are commonly used in the literature covering deal complexity as variables to orthogonalize our textual disclosure quality and quantity measures. For this, we again follow Zhang et al. (2023) and regress each measure on *Number of securities* and *Rating disagreement* as proxies for structural complexity and on *Number of loans*, the number of loans in the ABS deal, and *SD loan interest rates*, the standard deviation of the interest rates of the loans in the ABS deal, as proxies for collateral complexity. We then use the residuals from each of these regressions and revisit the central findings of our previous analyses.

The results can be found in Tables 8 and 9. For the direct effect of disclosure quality and quantity on *Yield spread* as well as with respect to the predictive ability of *Yield spread*, we confirm our previous findings as shown above. Therefore, we follow that our results are not driven by the complexity of the underlying ABS deals, suggesting that it is indeed the textual disclosure quality and quantity affecting investors' security pricing.

[Tables 8 and 9 about here.]

V.5 Textual disclosure quality and quantity in the risk factor section of prospectuses

After having examined measures of textual disclosure quality and quantity based on the texts of the entire ABS prospectuses, in the next step of our empirical analysis, we focus on a specific section of ABS prospectuses, namely the risk factor section. As described in Section II.1, since the implementation of the EU Prospectus Directive in 2004, it is mandatory to include a section in prospectuses that describes the risk factors of the respective security. This analysis is also motivated by recent studies highlighting the importance of the risk factor section in ABS prospectuses and other security-related disclosures (e.g., Cazier et al., 2021; Zhang et al., 2023). Thus, we conjecture that the risk factor section should be particularly important for the risk assessment of investors. To examine the risk factor section in ABS prospectuses, we re-estimate the regression models from Equations

V.1 and V.2 but now specifically measure the textual disclosure quality and quantity only of the section in the ABS prospectuses where risk factors are described.[20] The results can be found in Tables 10 and 11.

For the direct effect of disclosure quality and quantity on *Yield spread*, we confirm our previous findings and find a negative and statistically significant coefficient for *Boilerplate language* and a positive and significant coefficient for *Risk factor length*. The coefficients for the interactions between *Boilerplate language*, *Linguistic complexity*, and *Risk factor length* on the one side and *Yield spread* on the other side remain negative and statistically significant. Overall, we infer that mostly more standardized descriptions of risk factors, on average, decrease the initial yield spread investors demand. This again supports the idea that common phrases describing risks are perceived as a positive signal by investors. Longer descriptions of risks, however, relate to a higher initial yield spread and likely signal higher risk to investors. The latter finding is further supported by the (unreported) result that the relative length of the risk factor section, calculated as the length of the risk factor section divided by the length of the entire prospectus, yields a positive and statistically significant coefficient. Thus, a larger share of the description of risks within the prospectuses is also seen as a signal of higher risk by investors.

[Tables 10 and 11 about here.]

V.6 Textual disclosure quality and quantity and the parties involved in the prospectus creation

We now shed more light on the creation of the prospectuses and the parties that are involved in the process. Most importantly, the originator is responsible for supplying the vast majority of information provided to the investors, especially for the description of

[20]For this analysis, we adjust the measures for textual disclosure quality and quantity accordingly. *Boilerplate language* is now calculated based on tetragrams that often appear in the risk factor section of ABS prospectuses related to the same asset class. *Linguistic complexity* is now calculated as the readability of the text in the risk factor section. Our new measure of disclosure quantity, *Risk factor length*, is now defined as the logarithmized number of words in the risk factor section.

the underlying loan portfolio and the risk factors related to the ABS. In our framework of the securitization market, we take advantage of the unique setting in the issuance process that information on bank loans is highly confidential, protected even by law, and the deal structure is individual to the respective deal. Consequently, investors have no way to receive information on the specific loan portfolio and deal structure prior to the information disclosure provided by the bank in the issuance prospectus.

As the second important party, typically, one or more law firms advise the issuer in the issuance process and thus are involved in preparing the prospectus. Thus, we manually collected these law firms involved from all prospectuses in our data set. On average, 1.8 law firms work on one issuance prospectus in our data set with a minimum number of one and a maximum of seven. As law firms are engaged in a large number of issuances, often distributed across different issuers, we have 138 different law firms in our sample. The important role of lawyers in sales processes is also highlighted by Karsten et al. (2021). Even though the focus of their study is in the context of M&A, it underlines the importance of providing more detailed insights into the impact of the involved lawyers on our results. In our framework, investigating the influence of the law firms is particularly important, as different law firms have different writing styles that, in a first step, influence the values of our textual disclosure quality and quantity measures. Thus, in a second step, the law firms' impact on the linguistic design of the prospectus could affect investors' risk assessment and could plausibly be correlated with the yield spreads demanded at issuance.

We add law firm fixed effects for those law firms that are involved in at least 20 deals and an additional dummy variable that is one for all those deals, in which law firms with less than 20 deal participations are involved, and zero otherwise. The results are presented in Tables 12 and 13 and predominantly confirm our results shown in Sections V.1 and V.2, while the coefficient of *Prospectus length* is only significant when including all three measures of textual disclosure quality and quantity jointly (column (4)). Overall, even though law firms take an important role in the prospectus creation process, our results are not driven by the individual peculiarities of these firms.

[Tables 12 and 13 about here.]

V.7 Textual disclosure quality and quantity and regulation

After providing evidence that the textual disclosure quality and quantity in ABS prospectuses and specifically in the risk factor section affect investors in their pricing and leads to less accurate risk assessments, we examine in the final step of our empirical analysis whether recent EU regulations have successfully strengthened investor protection by leading to a reduction in *Boilerplate language*, *Linguistic complexity*, and *Prospectus length*. Specifically, we investigate the EU Prospectus Regulation and the Securitization framework described in Section II.1, which were introduced in 2017 and intended to increase the informativeness of prospectuses by demanding a concise (relates to prospectus length) and comprehensible (relates to linguistic complexity) communication of material and specific (relates to boilerplate language) information. In this way, we also provide evidence of the implications of regulatory actions on textual disclosures, as suggested by Leuz and Wysocki (2016).

Empirical strategy:

To examine whether and how the EU Prospectus Regulation and the Securitization framework affected ABS prospectuses, we use an empirical strategy based on a difference-in-differences approach that is similar to the approach used by deHaan et al. (2021). As the dependent variable, we use each of the three textual disclosure measures individually. As independent variables, we use an indicator variable *Post*, which is one for the years 2018 and after. This variable marks the years after the adoption of the EU Prospectus Regulation and the Securitization framework and is zero for all years up to 2016.[21] As 2017 itself is characterized by major regulatory changes during the year, this year is removed from the analysis so that a clear allocation of all observations to the post- or pre-period is possible. For this analysis, we only include the years 2010 and after to ensure a stable

[21] Although both regulations took effect in 2019, they were adopted in 2017, and hence, issuers were aware of it from this year on.

regulatory environment prior to the shock.[22] However, our results remain unchanged if we include the years 2002 to 2009 in our analysis sample (see Table A.10 in the Internet Appendix).

As a second independent variable, we further use a treatment variable for each of the textual disclosure measures (*Boilerplate language (high)*, *Prospectus length (high)*, *Linguistic complexity (high)*), which equals one if the mean value of the respective measure for the issuer of the ABS is above the 60% quantile of the measure before the regulations and zero if the mean value is below the 40% quantile of the measure before the regulations. Thereby, the variable indicates whether the issuer used particularly standardized, complex, and long prospectuses before the regulations or not. Most importantly for this analysis, we include an interaction term between the respective treatment variable and *Post* to identify, whether the EU regulations affected the textual disclosure quality and quantity. We further control for deal origination year FE, asset class FE, and country FE.[23] We run the analysis for the textual disclosure quality and quantity measures based on the entire prospectus and based solely on the risk factor section because the risk factor section received particular attention in the EU Prospectus Regulation (see Section II.1 for further details). Specifically, we run the following regression:

$$\begin{aligned} &Text.\ discl.\ quality\ and\ quantity_i \\ &\quad = \beta_0 + \beta_1\ x\ Text.\ discl.\ quality\ and\ quantity\ (high)_i \\ &\qquad + \beta_2\ x\ Text.\ discl.\ quality\ and\ quantity\ (high)_i\ \text{x}\ Post \\ &\qquad + \alpha_i + \gamma_i + \zeta_i + \epsilon_i \end{aligned} \tag{V.3}$$

As before, we estimate the regression coefficients using an OLS estimator. In this analysis, we cluster standard errors at the bank level. To assess the validity of the empirical strategy, we examine the assumption of parallel trends in Figure A.7 in the Internet Appendix. Overall, we observe reasonably parallel trends for our variables of interest.

[22] As discussed in Section II.1, the last important regulatory change before the 2017 regulations was the Amending Directive in 2010.

[23] Note that the coefficient for *Post* is absorbed by the deal origination year FE.

Results:

The results of this analysis are presented in Table 14. For the textual disclosure quality and quantity measures based on the entire prospectus, we see that the coefficients for each of the measures are positive and statistically significant. As expected, it indicates that issuers writing prospectuses with high levels of *Boilerplate language*, *Linguistic complexity*, and *Prospectus length* before the regulations constantly exhibit more standardized, complexer, and lengthier descriptions over the whole sample period, that is, before and also after the regulatory intervention. Most importantly, the coefficients of the interaction terms for *Boilerplate language (high)* and *Prospectus length (high)* on the one side and *Post* on the other side are negative and statistically significant. This indicates that issuers with particularly long and standardized prospectuses prior to 2017 drafted shorter and more individual prospectuses after the announcement of the EU regulations in 2017 compared to those issuers that provided less standardized and shorter prospectuses prior to 2017, compared to their average levels of our textual disclosure measures over the whole observation period. The coefficient of the interaction term for *Linguistic complexity* and *Post*, however, is negative yet not statistically significant.

[Table 14 about here.]

When examining the results for the textual disclosure quality and quantity measures based on the risk factor section of the prospectuses, the previous results are strongly supported. We again find negative and statistically significant coefficients of the interaction terms for *Boilerplate language (high)* and *Prospectus length (high)* on the one side and *Post* on the other side. Interestingly, now the coefficient of the interaction term for *Linguistic complexity* and *Post* is also negative and statistically significant. We conclude that the impact of the 2017 regulations is more far-reaching for the risk factors sections of ABS prospectuses than for the prospectuses as a whole.

When interpreting the results for the effect of the EU regulations in 2017 in Table 14 in combination with Figure A.7 in the Internet Appendix, we see that boilerplate language

indeed decreased for the treatment group, while for linguistic complexity and prospectus length, this only holds for the risk factor section. We also see that the average values of some measures for the control group tend to increase after the regulation, especially when calculating the measures based on the entire prospectus. This is remarkable as the new regulations themselves should not lead issuers to make prospectuses more standardized, complex, or longer. Nevertheless, some of those issuers in our sample that have less complex and shorter prospectuses prior to 2017 decreased their textual disclosure quality and increased the quantity in the aftermath of the regulation. Overall, these results indicate that the regulatory actions homogenized the textual disclosure quality and quantity in the European ABS market, as all of the three textual disclosure measures tend to be closer to their respective average in our entire sample.

VI Conclusion

Understanding the characteristics and thus the risk of a security is central to investors' pricing of the security. Prospectuses contain important information on securities and specifically on their risk. This is particularly so for asset-backed securities (ABS) issuances, where investors have no publicly available information on important security characteristics, such as the quality of the underlying loan portfolio or the deal structure, and consequently have to rely on the information provided in the prospectus. Recent evidence documents that textual information in ABS prospectuses is highly relevant but not properly priced by investors (Zhang et al., 2023). In this context, we examine how the quality and quantity of textual disclosure in ABS prospectuses affect investors' security pricing.

Building on a unique sample of all ABS deals in Europe reported under the loan-level initiative of the European Central Bank supplemented by market data of S&P Global, we find that the share of boilerplate language, the linguistic complexity, and the length of ABS prospectuses substantially affect investors' security pricing at issuance beyond all observable risk factors. Furthermore, the affected prices are less informative of future

security performance. Our findings add to a small but growing literature examining the impact of prospectuses on investors and shed light on the role of information provision for security pricing and, in particular, ABS pricing by showing that the quality and quantity of textual information are important determinants of demanded yields. More generally, our results provide important insights for the scientific debate on the impact of textual disclosures on the perception of risk and the resulting market outcomes.

Our findings have important implications for issuers, investors, and regulators alike as they provide novel insights into how investors price securities and why they struggle to adequately assess risk, which is critical for efficient capital markets. Specifically, issuers should be aware of the consequences of their prospectus design, as the textual disclosure quality and quantity can turn out to their disadvantage in the prices of their issued securities. Investors should try to let their pricing stay unaffected by the prospectus design. To support investors in adequate security pricing, regulators should continue to pay significant attention to the textual disclosure in security issuance prospectuses.

References

Abis, S., A. M. Buffa, A. Javadekar, and A. Lines (2021). Learning from prospectuses. Working Paper. Columbia Business School. December 24, 2021.

Association for Financial Markets in Europe (2014). AFME securitisation data report, fourth quarter 2014. https://www.sifma.org/resources/research/afme-securitisation-data-report-fourth-quarter-2014/ (Last accessed: 20.05.2021).

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

Becker, B. and T. Milbourn (2011). How did increased competition affect credit ratings? *Journal of Financial Economics 101*(3), 493–514.

Billio, M., A. Dufour, S. Segato, and S. Varotto (2023). Complexity and the default risk of mortgage-backed securities. *Journal of Banking & Finance*, 106993.

Blankespoor, E., E. deHaan, and I. Marinovic (2020). Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics 70*(2), 101344.

Bloomfield, R. (2008). Discussion of "Annual report readability, current earnings, and earnings persistence". *Journal of Accounting and Economics 45*(2-3), 248–252.

Bonsall, S. B., A. J. Leone, B. P. Miller, and K. Rennekamp (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics 63*(2), 329–357.

Bonsall, S. B. and B. P. Miller (2017). The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies 22*(2), 608–643.

Bourveau, T., E. T. De George, A. Ellahie, and D. Macciocchi (2022). The role of disclosure and information intermediaries in an unregulated capital market: Evidence from initial coin offerings. *Journal of Accounting Research 60*(1), 129–167.

Bushee, B. J., I. D. Gow, and D. J. Taylor (2018). Linguistic complexity in firm disclosures: Obfuscation or information? *Journal of Accounting Research 56*(1), 85–121.

Cai, K. N., K. W. Hanley, A. G. Huang, and X. Zhao (2022). Risk disclosure and the pricing of corporate debt issues in private and public markets. Working Paper. *Georgetown University McDonough School of Business Research Paper No. 3700162.*

Cazier, R. A., J. L. McMullin, and J. S. Treu (2021). Are lengthy and boilerplate risk factor disclosures inadequate? An examination of judicial and regulatory assessments of risk factor language. *The Accounting Review 96*(4), 131–155.

Celérier, C. and B. Vallée (2017). Catering to investors through security design: Headline rate and complexity. *The Quarterly Journal of Economics 132*(3), 1469–1508.

Chen, W., P. Hribar, and S. Melessa (2018). Incorrect inferences when using residuals as dependent variables. *Journal of Accounting Research 56*(3), 751–796.

Chy, M. and H. Kyung (2023). The effect of bond market transparency on bank loan contracting. *Journal of Accounting and Economics 75*(2), 101536.

Cox, C. (2007). Speech by SEC chairman: Closing remarks to the second annual corporate governance summit. USC Marshall School of Business.

deHaan, E., Y. Song, C. Xie, and C. Zhu (2021). Obfuscation in mutual funds. *Journal of Accounting and Economics 72*(2-3), 101429.

Dyer, T., M. Lang, and L. Stice-Lawrence (2017). The evolution of 10-K textual disclosure: Evidence from latent dirichlet allocation. *Journal of Accounting and Economics 64*(2-3), 221–245.

Easley, D. and M. O'Hara (2004). Information and the cost of capital. *The Journal of Finance 59*(4), 1553–1583.

Ehrmann, M. and J. Talmi (2020). Starting from a blank page? Semantic similarity in central bank communication and market volatility. *Journal of Monetary Economics 111*, 48–62.

Ellison, G. and S. F. Ellison (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica 77*(2), 427–452.

Ellison, G. and A. Wolitzky (2012). A search cost model of obfuscation. *The RAND Journal of Economics 43*(3), 417–441.

Enriques, L. and M. Gatti (2008). Is there a uniform EU securities law after the Financial Services Action Plan? *Stanford Journal of Law, Business & Finance 14*, 43–81.

Ertan, A., M. Loumioti, and R. Wittenberg-Moerman (2017). Enhancing loan quality through transparency: Evidence from the European Central Bank loan level reporting initiative. *Journal of Accounting Research 55*, 877–918.

Ertugrul, M., J. Lei, J. Qiu, and C. Wan (2017). Annual report readability, tone ambiguity, and the cost of borrowing. *Journal of Financial and Quantitative Analysis 52*(2), 811–836.

European Commission (2015). Review of the prospectus directive. Consultation document.

European Securities and Markets Authority (2022). Parsing prospectuses: A text-mining approach. ESMA Report on Trends, Risks and Vulnerabilities Risk Analysis.

European Systemic Risk Board (2022). Monitoring systemic risks in the EU securitisation market. ESRB report July 2022.

Fischer-Appelt, D. (2010). Prospectus directive amendments—discussion of key changes. *Law and Financial Markets Review 4*(5), 490–498.

Furfine, C. H. (2014). Complexity and loan performance: Evidence from the securitization of commercial mortgages. *The Review of Corporate Finance Studies 2*(2), 154–187.

Ghent, A., W. N. Torous, and R. Valkanov (2019). Complexity in structured finance. *The Review of Economic Studies 86*(2), 694–722.

Hanley, K. W. and G. Hoberg (2010). The information content of IPO prospectuses. *The Review of Financial Studies 23*(7), 2821–2864.

Hansen, S., M. McMahon, and A. Prat (2017). Transparency and deliberation within the FOMC: A computational linguistics approach. *The Quarterly Journal of Economics 133*(2), 801–870.

He, J., J. Qian, and P. E. Strahan (2016). Does the market understand rating shopping? Predicting MBS losses with initial yields. *The Review of Financial Studies 29*(2), 457–485.

Helwege, J. and C. M. Turner (1999). The slope of the credit yield curve for speculative-grade issuers. *The Journal of Finance 54*, 1869–1884.

Hibbeln, M. T., R. Metzler, and W. Osterkamp (2022). Not on the same page — (text-) complexity in European securitizations. Working Paper. Universität Duisburg-Essen - Mercator School of Management. August 5, 2022.

Hibbeln, M. T. and W. Osterkamp (2020). The impact of skin in the game on bank behavior in the securitization market. Working Paper. Universität Duisburg-Essen - Mercator School of Management. January 17, 2020.

Howell, E. (2018). An analysis of the prospectus regime: The EU reforms and the 'Brexit' factor. *European Company and Financial Law Review 15*(1), 69–100.

Karsten, C., U. Malmendier, and Z. Sautner (2021). Lawyer expertise and contract design — Evidence from M&A negotiations. *The Economic Journal 132*(642), 644–674.

Klein, P., C. Mössinger, and A. Pfingsten (2021). Transparency as a remedy for agency problems in securitization? The case of ECB's loan-level reporting initiative. *Journal of Financial Intermediation 46*, 100853.

Kleymenova, A. and R. E. Tomy (2022). Observing enforcement: Evidence from banking. *Journal of Accounting Research 60*(4), 1583–1633.

Kravet, T. and V. Muslu (2013). Textual risk disclosures and investors' risk perceptions. *Review of Accounting Studies 18*(4), 1088–1122.

Lang, M. and L. Stice-Lawrence (2015). Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics 60*(2-3), 110–135.

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics 56*(1), 130–147.

Leuz, C. and P. D. Wysocki (2016). The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research. *Journal of Accounting Research 54*(2), 525–622.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics 45*(2-3), 221–247.

Lo, K., F. Ramos, and R. Rogo (2017). Earnings management and annual report readability. *Journal of Accounting and Economics 63*(1), 1–25.

Loughran, T. and B. McDonald (2014). Measuring readability in financial disclosures. *The Journal of Finance 69*(4), 1643–1671.

Lurie, N. H. and C. H. Mason (2007). Visual representation: Implications for decision making. *Journal of Marketing 71*(1), 160–177.

Mählmann, T. (2012). Did investors outsource their risk analysis to rating agencies? Evidence from ABS-CDOs. *Journal of Banking & Finance 36*(5), 1478–1491.

McClane, J. (2019). Boilerplate and the impact of disclosure in securities dealmaking. *Vanderbilt Law Review 72*(1), 191–295.

Meade, E. E. and D. Stasavage (2008). Publicity of debate and the incentive to dissent: Evidence from the US Federal Reserve. *The Economic Journal 118*(528), 695–717.

Miller, B. P. (2010). The effects of reporting complexity on small and large investor trading. *The Accounting Review 85*(6), 2107–2143.

Myers, S. C. and N. S. Majluf (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics 13*(2), 187–221.

Neilson, J., S. Ryan, P. Wang, and B. Xie (2022). Asset-level transparency and the (e)valuation of asset-backed securities. *Journal of Accounting Research 60*(3), 1131–1183.

Peña-Cerezo, M. Á., A. Rodríguez-Castellanos, and F. J. Ibáñez-Hernández (2019). Multi-tranche securitisation structures: More than just a zero-sum game? *The European Journal of Finance 25*(2), 167–189.

Seo, H. (2021). Peer effects in corporate disclosure decisions. *Journal of Accounting and Economics 71*(1), 101364.

Umar, T. (2022). Complexity aversion when seeking alpha. *Journal of Accounting and Economics 73*(2), 101477.

Zhang, H. H., F. Zhao, and X. Zhao (2023). Neglected risks in the communication of residential mortgage-backed securities offering. *The Journal of Finance, forthcoming*.

Zippia, Inc. (2023). What degree do you need to be an investment banker?, URL https://www.zippia.com/investment-banker-jobs/education (Last accessed: 10/5/2023).

VII Appendix

Table 1: Summary statistics

Variable	N	Mean	SD	Min	Median	Max
Textual disclosure quality and quantity (security level)						
Boilerplate language	2,469	0.39	0.20	0.03	0.44	0.78
Linguistic complexity	2,469	23.73	1.33	20.50	23.85	26.21
Prospectus length	2,469	11.42	0.37	10.57	11.48	12.09
Textual disclosure quality and quantity (deal level)						
Boilerplate language	1,014	0.43	0.19	0.03	0.45	0.78
Linguistic complexity	1,014	23.95	1.31	20.50	24.09	26.21
Prospectus length	1,014	11.43	0.33	10.57	11.47	12.09
Ex post performance						
Interest shortfall	2,469	0.45	5.01	0.00	0.00	100.00
Investors' pricing						
Yield spread	2,469	0.91	1.14	-1.93	0.60	6.75
Controls						
Interest rate	2,469	3.67	1.92	0.00	3.21	11.82
Number of securities	2,469	6.46	9.23	1	5	83
Rating disagreement	2,469	0.52	0.50	0	1	1
Security width	2,469	31.79	36.33	0.06	9.67	100.00
Security size	2,469	18.59	1.85	14.22	18.66	22.31
Security term	2,469	3.35	0.63	1.18	3.60	4.48
Excess interest	2,469	3.37	9.01	0.00	0.13	70.90
Subordination	2,469	17.41	22.13	0.00	8.28	92.11

This table reports the descriptive statistics for the variables used in our empirical analysis. Variables are described in Table A.2 in the Appendix. "N" denotes the number of non-missing values, "Mean" the mean, "SD" the standard deviation, "Min" the minimum, "Median" the median, and "Max" the maximum.

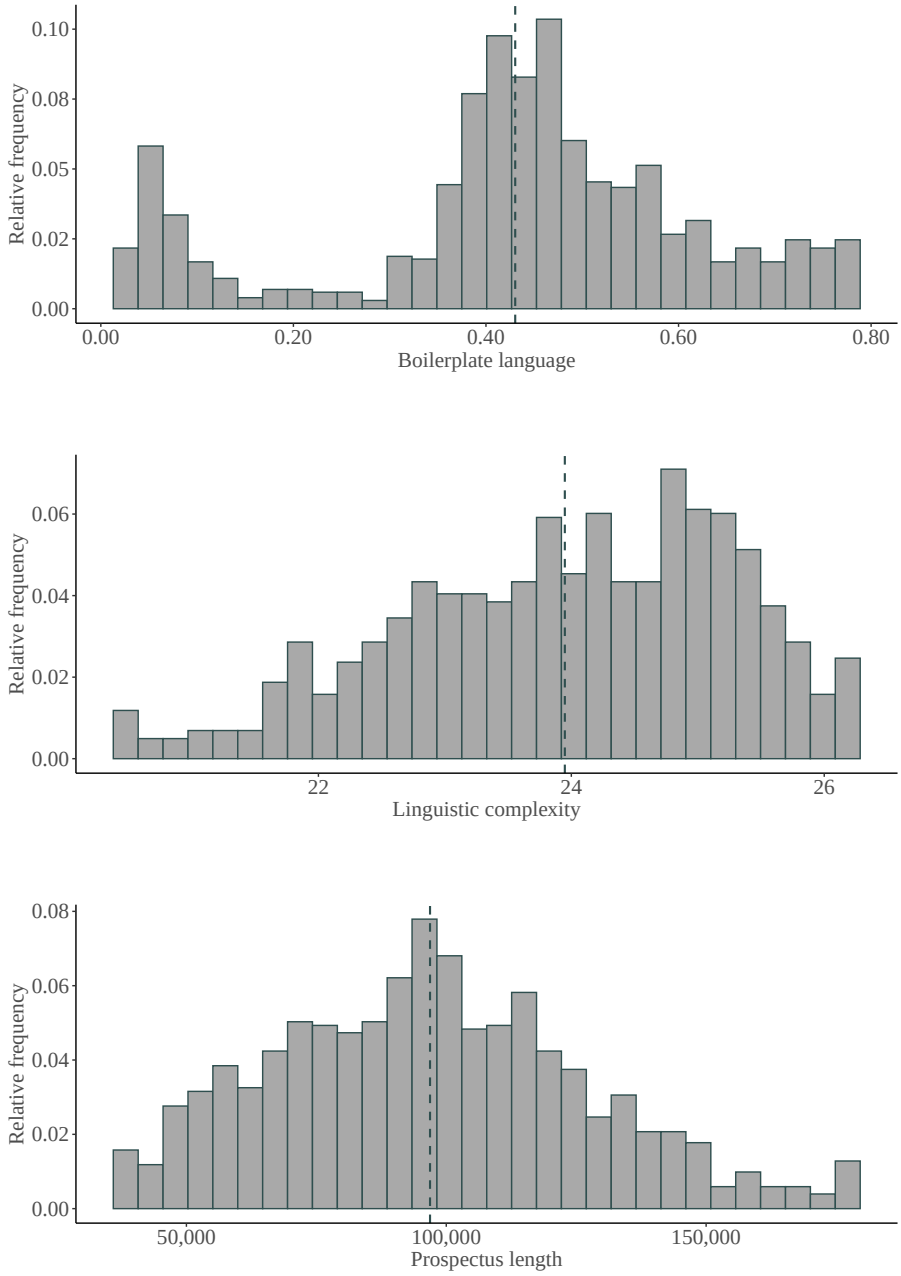

Figure 1: Distribution of textual disclosure quality and quantity measures

This figure displays the distribution of the share of boilerplate language, as measured by the occurrence of common tetragrams, the linguistic complexity, as measured by Gunning's Fog index, and the length of the prospectus, as measured by the number of words, in ABS prospectuses. Dashed vertical lines indicate the sample mean.

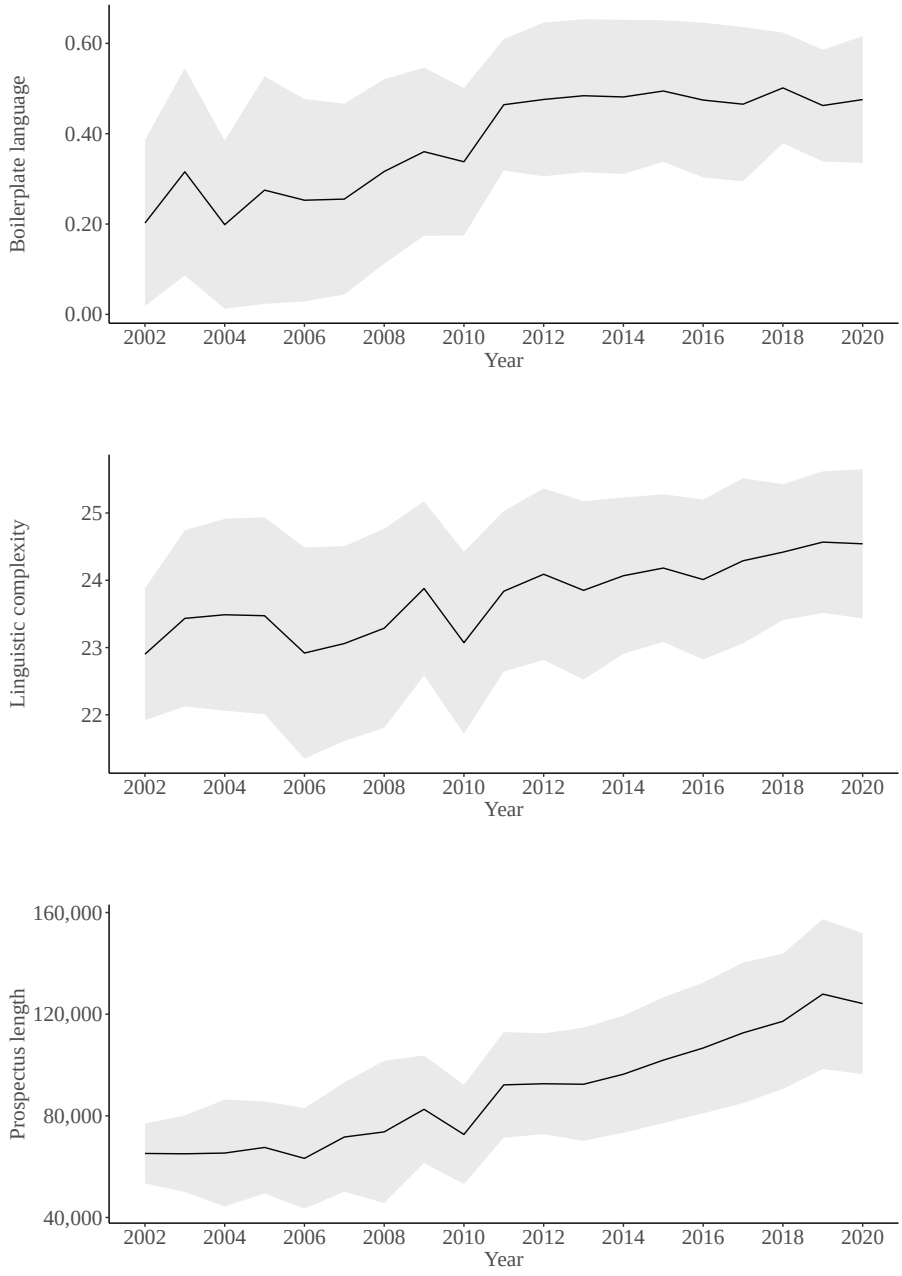

Figure 2: Development of textual disclosure quality and quantity measures over time

This figure displays the aggregated development of the share of boilerplate language, as measured by the occurrence of common tetragrams, the linguistic complexity, as measured by Gunning's Fog index, and the length of the prospectus, as measured by the number of words, in ABS prospectuses over time as well as intervals of one standard deviation above and below the mean of the measures per year.

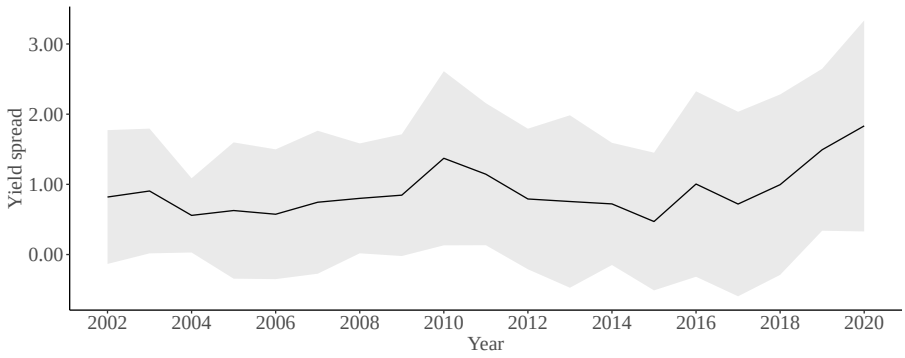

Figure 3: Development of the yield spread over time

This figure displays the aggregated development of the demanded yield spread for the ABS over time as well as intervals of one standard deviation above and below the mean per year.

Table 2: Textual disclosure quality and quantity and the yield spread

	Yield spread			
	(1)	(2)	(3)	(4)
Boilerplate language	-0.830***			-0.769**
	(0.315)			(0.306)
Linguistic complexity		-0.029		-0.026
		(0.025)		(0.024)
Prospectus length			0.311**	0.337***
			(0.125)	(0.124)
Interest rate	0.060***	0.058***	0.062***	0.059***
	(0.017)	(0.017)	(0.017)	(0.017)
Number of securities	0.003	0.002	0.003*	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Rating disagreement	0.168***	0.153***	0.160***	0.157***
	(0.048)	(0.049)	(0.048)	(0.048)
Security width	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Security size	-0.111***	-0.104***	-0.113***	-0.102***
	(0.022)	(0.022)	(0.022)	(0.022)
Security term	-0.075	-0.091	-0.069	-0.101
	(0.118)	(0.116)	(0.118)	(0.117)
Excess interest	0.013***	0.013***	0.013***	0.014***
	(0.004)	(0.004)	(0.004)	(0.004)
Subordination	-0.002	-0.002*	-0.002	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,469	2,469	2,469	2,469
Adj. R^2	0.60	0.60	0.60	0.61

This table displays the results of the analysis how measures of textual disclosure quality and quantity in ABS prospectuses are related to the initial yield spread demanded by investors (*Yield spread*). The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 3: Textual disclosure quality and quantity and the yield spread: Mechanisms and insights

	Yield spread								
Mechanism	*Rating*	*Deal count*	*Visual.*	*Rating*	*Deal count*	*Visual.*	*Rating*	*Deal count*	*Visual.*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Boilerplate language x	0.0956***	0.0260**	-0.644***						
	(0.0260)	(0.0102)	(0.2259)						
Linguistic complexity x				0.00864**	0.00317**	-0.0753**			
				(0.0040)	(0.0016)	(0.0332)			
Prospectus length x							0.0554***	0.0346***	-0.327***
							(0.0165)	(0.0073)	(0.1192)
Textual discl. measure	-1.298***	-1.263***	-0.427	-0.0737***	-0.0590	0.00610	-0.00739	-0.071	0.476***
	(0.3155)	(0.3548)	(0.3152)	(0.0253)	(0.0380)	(0.0249)	(0.1248)	(0.1503)	(0.1250)
Interaction variable	-	-0.0012	0.216**	-	-0.067*	1.729**	-	-0.386***	3.641***
		(0.0031)	(0.1070)		(0.0361)	(0.7939)		(0.0813)	(1.3618)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,469	2,469	2,435	2,469	2,469	2,435	2,469	2,469	2,435
Adj. R^2	0.61	0.61	0.61	0.60	0.60	0.61	0.61	0.61	0.61

The table displays the results of the analysis how measures of disclosure quality and quantity in ABS prospectuses are related to the initial yield spread demanded by investors (*Yield spread*) providing evidence on three different mechanisms determining this relation. Variables are described in Table A.2. The table reports standard errors clustered with respect to the ABS deal in parentheses. *p<0.1; **p<0.05; ***p<0.01.

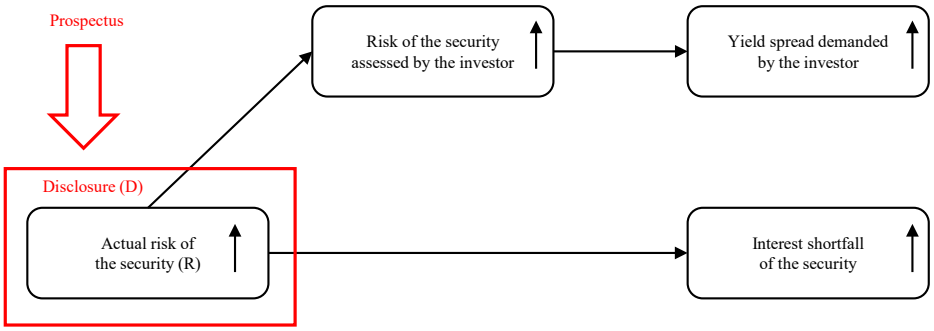

Figure 4: Predictive ability of the yield spread and the impact of disclosure

This figure displays the mechanism of how the initial yield spread demanded by investors predicts security performance as measured by the interest shortfall and the role of the textual disclosure in the emission prospectuses on this relation.

Table 4: Textual disclosure quality and quantity and the predictive ability of the yield spread

	Interest shortfall			
	(1)	(2)	(3)	(4)
Yield spread	3.149***	10.280***	25.880***	7.443*
	(0.864)	(3.408)	(7.168)	(4.406)
Boilerplate language x Yield spread	-6.359***			-6.036***
	(1.651)			(1.821)
Linguistic complexity x Yield spread		-0.414***		0.115
		(0.141)		(0.213)
Prospectus length x Yield spread			-2.212***	-0.622
			(0.611)	(0.629)
Boilerplate language	5.443***			5.282***
	(1.517)			(1.581)
Linguistic complexity		0.465***		-0.069
		(0.135)		(0.104)
Prospectus length			3.111***	1.714***
			(0.857)	(0.614)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,469	2,469	2,469	2,469
Adj. R^2	0.23	0.20	0.21	0.23

This table displays the results of the analysis how the interaction between measures of textual disclosure quality and quantity in ABS prospectuses on the one hand and the initial yield spread demanded by investors (*Yield spread*) on the other hand is related to the ex post performance of ABS (*Interest shortfall*), indicating the predictive ability of yield spreads at security issuance. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

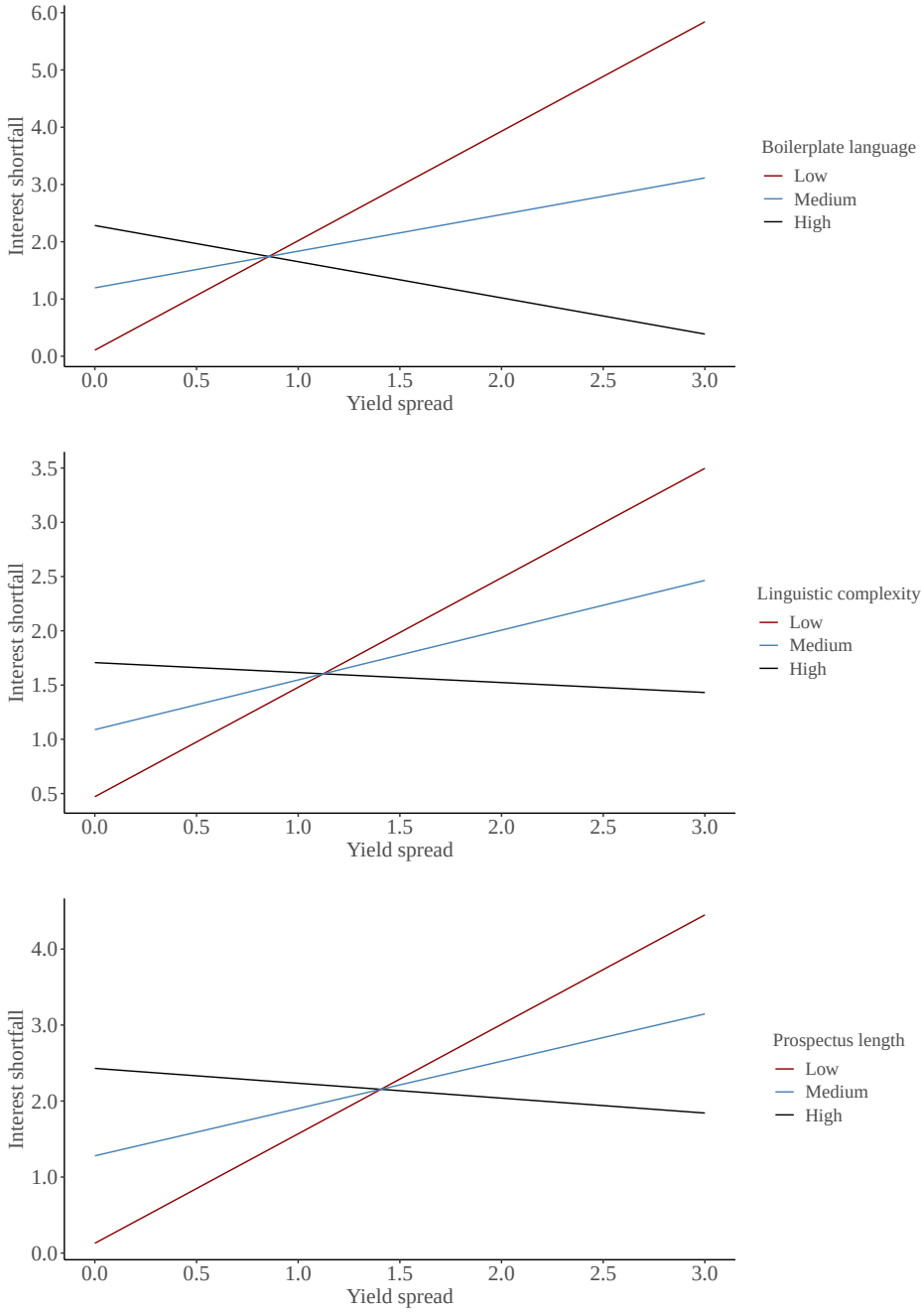

Figure 5: Yield spread, interest shortfall, and textual disclosure quality and quantity

This figure displays the relationship between the initial yield spread demanded by investors (*Yield spread*) and the ex post performance of ABS (*Interest shortfall*) at different levels of our textual disclosure quality and quantity measures.

Table 5: Textual disclosure quality and quantity and the secondary market spread volatility

	Spread volatility			
	(1)	(2)	(3)	(4)
Boilerplate language	0.193			0.149
	(0.139)			(0.144)
Linguistic complexity		0.022**		0.016
		(0.011)		(0.012)
Prospectus length			0.103**	0.089*
			(0.050)	(0.050)
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,468	2,468	2,468	2,468
Adj. R^2	0.40	0.40	0.40	0.40

This table displays the results of the analysis how measures of textual disclosure quality and quantity in ABS prospectuses are related to the volatility of the secondary market spread (*Spread volatility*). The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 6: Textual disclosure quality and quantity and the yield spread: ABS performance

	Yield spread			
	(1)	(2)	(3)	(4)
Boilerplate language	-0.883***			-0.818***
	(0.314)			(0.306)
Linguistic complexity		-0.0321		-0.0271
		(0.0249)		(0.0241)
Prospectus length			0.289**	0.315**
			(0.124)	(0.123)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,469	2,469	2,469	2,469
Adj. R^2	0.60	0.60	0.60	0.61

This table displays the results of the analysis how orthogonalized (regarding security risk and performance) measures of textual disclosure quality and quantity in ABS prospectuses are related to the initial yield spread demanded by investors (*Yield spread*). The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 7: Textual disclosure quality and quantity and the predictive ability of the yield spread: ABS performance

	Interest shortfall			
	(1)	(2)	(3)	(4)
Yield spread	0.500**	0.520**	0.373*	0.527**
	(0.211)	(0.219)	(0.205)	(0.247)
Boilerplate language x Yield spread	-7.251***			-6.507***
	(1.829)			(1.862)
Linguistic complexity x Yield spread		-0.529***		0.0970
		(0.158)		(0.208)
Prospectus length x Yield spread			-2.670***	-0.835
			(0.730)	(0.643)
Boilerplate language	2.558			2.390
	(1.772)			(1.694)
Linguistic complexity		0.362**		-0.126
		(0.144)		(0.111)
Prospectus length			2.024***	0.560
			(0.736)	(0.430)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,469	2,469	2,469	2,469
Adj. R^2	0.25	0.22	0.21	0.25

This table displays the results of the analysis how the interaction between orthogonalized (regarding security risk and performance) measures of textual disclosure quality and quantity in ABS prospectuses on the one hand and the initial yield spread demanded by investors (*Yield spread*) on the other hand is related to the ex post performance of ABS (*Interest shortfall*), indicating the predictive ability of yield spreads at security issuance. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 8: Textual disclosure quality and quantity and the yield spread: ABS complexity

	Yield spread			
	(1)	(2)	(3)	(4)
Boilerplate language	-0.935***			-0.908***
	(0.264)			(0.264)
Linguistic complexity		-0.038		-0.024
		(0.024)		(0.024)
Prospectus length			0.218*	0.280**
			(0.125)	(0.124)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,469	2,469	2,469	2,469
Adj. R^2	0.60	0.60	0.60	0.61

This table displays the results of the analysis how orthogonalized (regarding deal complexity) measures of textual disclosure quality and quantity in ABS prospectuses are related to the initial yield spread demanded by investors (*Yield spread*). The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 9: Textual disclosure quality and quantity and the predictive ability of the yield spread: ABS complexity

	Interest shortfall			
	(1)	(2)	(3)	(4)
Yield spread	0.525**	0.443*	0.561**	0.527**
	(0.227)	(0.232)	(0.234)	(0.266)
Boilerplate language x Yield spread	-5.160***			-4.884***
	(1.493)			(1.592)
Linguistic complexity x Yield spread		-0.359**		0.029
		(0.149)		(0.215)
Prospectus length x Yield spread			-1.788***	-0.276
			(0.569)	(0.636)
Boilerplate language	4.870***			4.600***
	(1.486)			(1.467)
Linguistic complexity		0.399***		-0.034
		(0.121)		(0.103)
Prospectus length			2.417***	1.190**
			(0.788)	(0.596)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,469	2,469	2,469	2,469
Adj. R^2	0.22	0.19	0.20	0.22

This table displays the results of the analysis how the interaction between orthogonalized (regarding deal complexity) measures of textual disclosure quality and quantity in ABS prospectuses on the one hand and the initial yield spread demanded by investors (*Yield spread*) on the other hand is related to the ex post performance of ABS (*Interest shortfall*), indicating the predictive ability of yield spreads at security issuance. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 10: Textual disclosure quality and quantity and the yield spread: Risk factors

	Yield spread			
	(1)	(2)	(3)	(4)
Boilerplate language	-0.847***			-0.727**
	(0.309)			(0.305)
Linguistic complexity		-0.012		-0.031*
		(0.019)		(0.018)
Risk factor length			0.299***	0.324***
			(0.078)	(0.081)
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,452	2,452	2,452	2,452
Adj. R^2	0.60	0.60	0.60	0.61

This table displays the results of the analysis how measures of textual disclosure quality and quantity in the risk factor section of ABS prospectuses are related to the initial yield spread demanded by investors (*Yield spread*). The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 11: Textual disclosure quality and quantity and the predictive ability of the yield spread: Risk factors

	Interest shortfall			
	(1)	(2)	(3)	(4)
Yield spread	1.602***	8.838***	13.99***	12.40***
	(0.474)	(3.045)	(3.538)	(3.881)
Boilerplate language x Yield spread	-5.943***			-3.677***
	(1.704)			(1.332)
Linguistic complexity x Yield spread		-0.333***		0.048
		(0.118)		(0.136)
Risk factor length x Yield spread			-1.394***	-1.285***
			(0.350)	(0.336)
Boilerplate language	4.947***			3.480***
	(1.550)			(1.243)
Linguistic complexity		0.309***		-0.013
		(0.109)		(0.103)
Risk factor length			1.231***	1.125***
			(0.390)	(0.388)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,452	2,452	2,452	2,452
Adj. R^2	0.21	0.20	0.25	0.25

This table displays the results of the analysis how the interaction between measures of textual disclosure quality and quantity in the risk factor section of ABS prospectuses on the one hand and the initial yield spread demanded by investors (*Yield spread*) on the other hand is related to the ex post performance of ABS (*Interest shortfall*), indicating the predictive ability of yield spreads at security issuance. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 12: Textual disclosure quality and quantity and the yield spread: Involved contract parties

	Yield spread			
	(1)	(2)	(3)	(4)
Boilerplate language	-0.800**			-0.794**
	(0.3880)			(0.3760)
Linguistic complexity		-0.0300		-0.0219
		(0.0275)		(0.0261)
Prospectus length			0.205	0.254*
			(0.1437)	(0.1427)
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Lawyer FE	Yes	Yes	Yes	Yes
Observations	2,343	2,343	2,343	2,343
Adj. R^2	0.62	0.62	0.62	0.62

This table displays the results of the analysis how measures of textual disclosure quality and quantity in ABS prospectuses are related to the initial yield spread demanded by investors (*Yield spread*) additionally considering the impact of the different parties involved in the prospectus creation. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 13: Textual disclosure quality and quantity and the predictive ability of the yield spread: Involved contract parties

	Interest shortfall			
	(1)	(2)	(3)	(4)
Yield spread	2.888***	10.47***	22.68***	5.985
	(0.8951)	(3.5122)	(6.6543)	(4.3161)
Boilerplate language x Yield spread	-5.875***			-5.518***
	(1.6943)			(1.8783)
Linguistic complexity x Yield spread		-0.425***		-0.00285
		(0.1446)		(0.1875)
Prospectus length x Yield spread			-1.939***	-0.278
			(0.5656)	(0.6055)
Boilerplate language	5.446***			4.685***
	(1.7281)			(1.7313)
Linguistic complexity		0.537***		0.0762
		(0.1637)		(0.1003)
Prospectus length			3.268***	1.762**
			(1.0071)	(0.7807)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Lawyer FE	Yes	Yes	Yes	Yes
Observations	2,343	2,343	2,343	2,343
Adj. R^2	0.22	0.19	0.20	0.22

This table displays the results of the analysis how the interaction between measures of textual disclosure quality and quantity in ABS prospectuses on the one hand and the initial yield spread demanded by investors (*Yield spread*) on the other hand is related to the ex post performance of ABS (*Interest shortfall*), indicating the predictive ability of yield spreads at security issuance, additionally considering the impact of the different parties involved in the prospectus creation. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 14: Textual disclosure quality and quantity and regulation

	Entire prospectus			Risk factor section		
	Boilerplate language	*Linguistic complexity*	*Prospectus length*	*Boilerplate language*	*Linguistic complexity*	*Prospectus length*
	(1)	(2)	(3)	(4)	(5)	(6)
Boilerplate language (high) x *Post*	-0.0452**			-0.0505***		
	(0.0194)			(0.0163)		
Linguistic complexity (high) x *Post*		-0.355			-0.810***	
		(0.2403)			(0.2947)	
Prospectus length (high) x *Post*			-0.120**			-0.382***
			(0.0475)			(0.1062)
Boilerplate language (high)	0.0757***			0.0367		
	(0.0177)			(0.0240)		
Linguistic complexity (high)		1.054***			1.347***	
		(0.3165)			(0.2711)	
Prospectus length (high)			0.222***			0.366***
			(0.0692)			(0.0997)
Deal orig. year FE	Yes	Yes	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	414	463	393	592	519	407
Adj. R^2	0.80	0.72	0.69	0.69	0.66	0.80

This table displays the results of the analysis how EU regulations in 2017 affected measures of textual disclosure quality and quantity in ABS prospectuses in general (columns (1) to (3)) and in the risk factor section (columns (4) to (6)). The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

A Internet Appendix

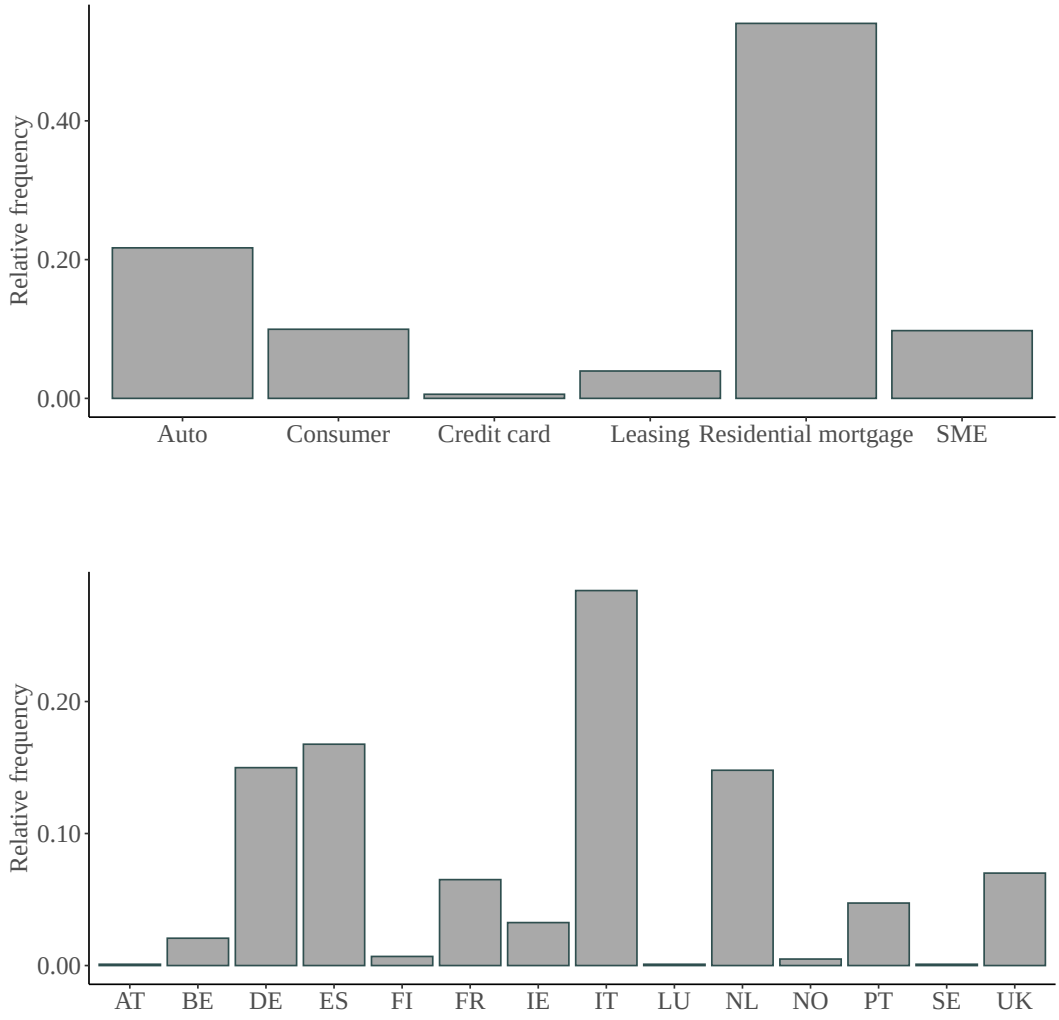

Figure A.1: Distribution of ABS deals

This figure displays the distribution of ABS deals across asset classes (upper plot) and across countries (bottom plot).

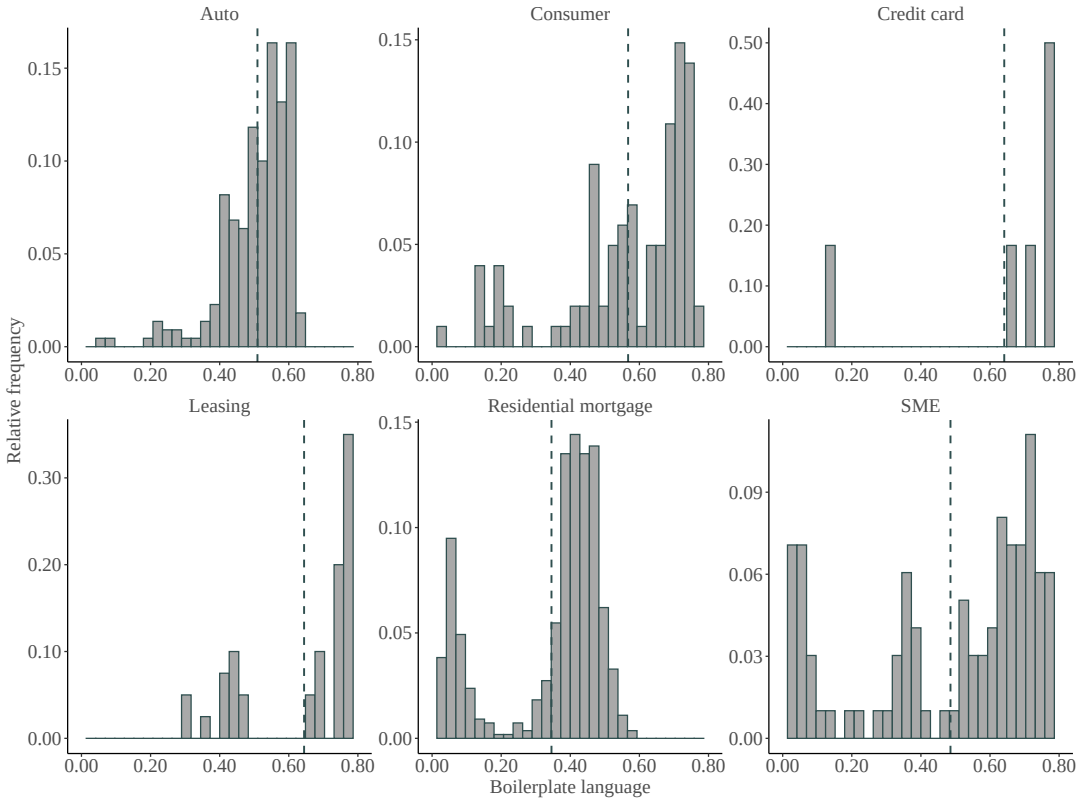

Figure A.2: Distribution of boilerplate language across asset classes

This figure displays the distribution of the share of boilerplate language, as measured by the occurrence of common tetragrams, in ABS prospectuses by asset class. Dashed vertical lines indicate the sample mean.

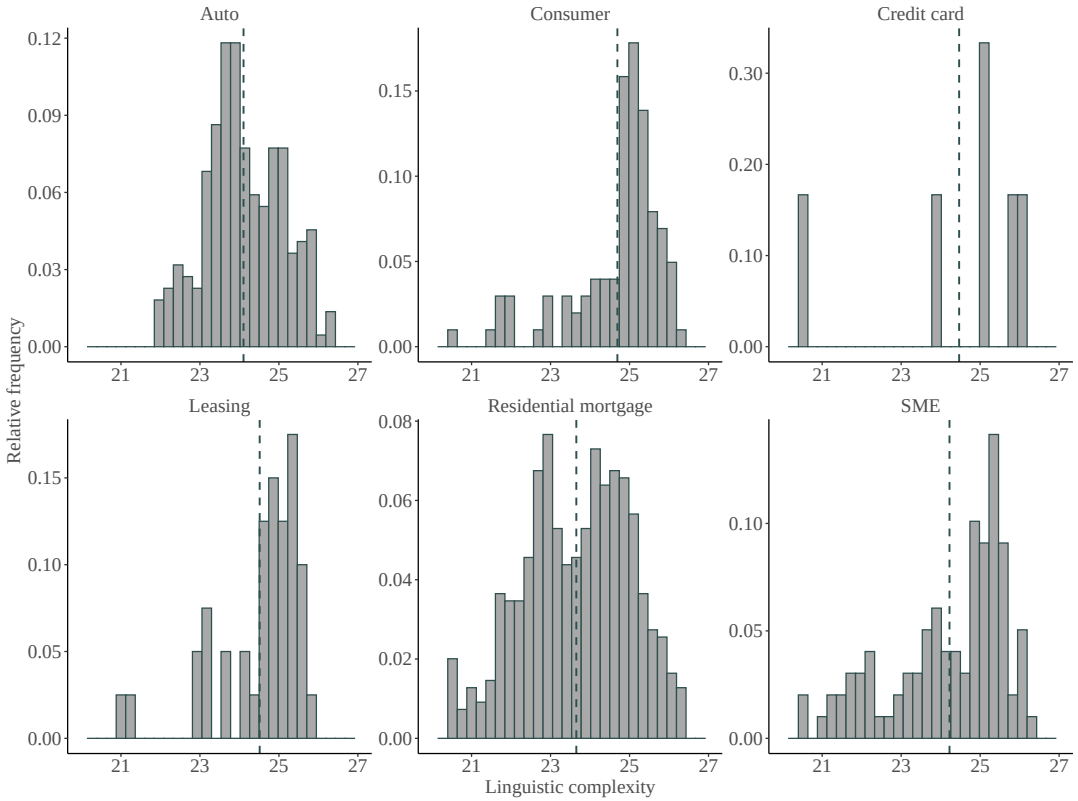

Figure A.3: Distribution of linguistic complexity across asset classes

This figure displays the distribution of the linguistic complexity, as measured by Gunning's Fog index, in ABS prospectuses by asset class. Dashed vertical lines indicate the sample mean.

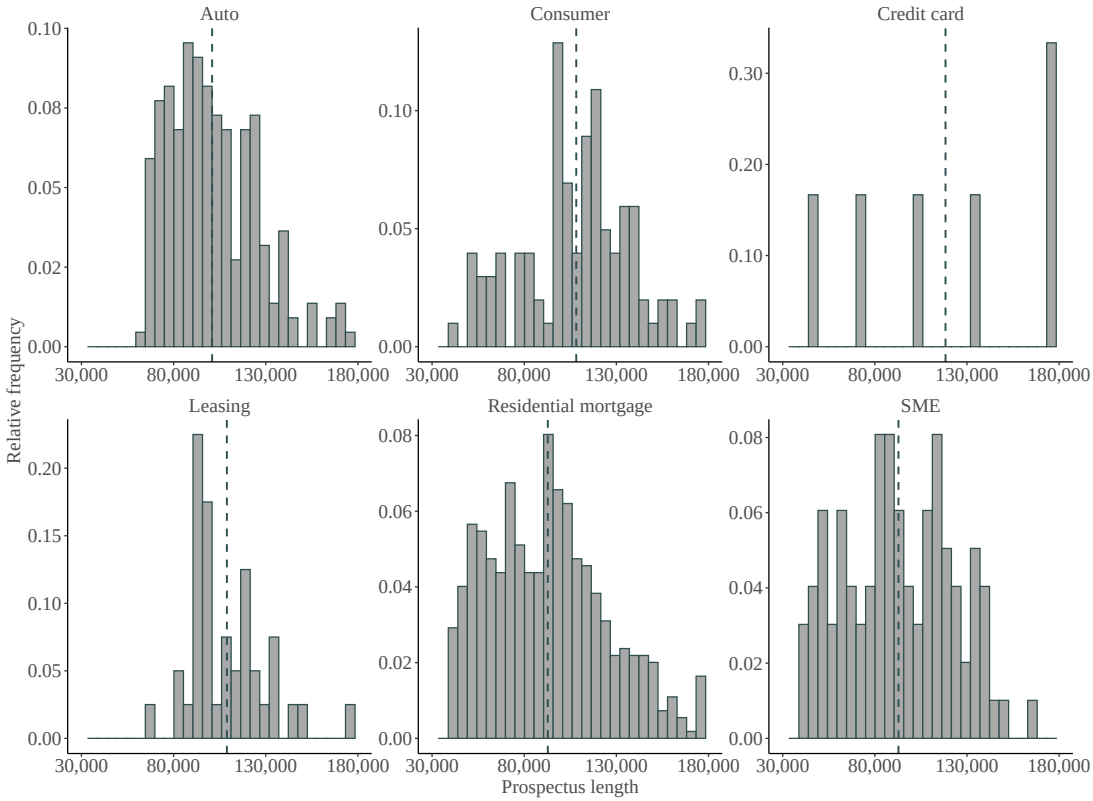

Figure A.4: Distribution of prospectus length across asset classes

This figure displays the distribution of the prospectus length, as measured by the number of words, in ABS prospectuses by asset class. Dashed vertical lines indicate the sample mean.

Prospectus A. Linguistic complexity (Gunning's Fog index): 19.52. Mean sentence length: 22.10.

The Receivables may be paid at any time and the Issuer cannot assure the Noteholders that new Receivables will be generated or will be generated at sufficient levels. To prevent the early redemption of the Notes, new Receivables must be generated. The Issuer is required to maintain a minimum amount of Receivables. The generation of new Receivables is affected by the Transferor's ability to compete in the then current industry environment and by customers' changing borrowing and payment patterns. If there is a decline in the generation of new Receivables or new accounts, the Noteholders may be repaid the principal before the expected date. In particular, the ability of the Issuer to utilise Partial Amortisation Amounts to make repayments of principal on the Notes to ensure the ongoing satisfaction of the Borrowing Base in the situation where the level of origination and therefore the amount of the Borrowing Base decreases.

No premium will be paid upon an early redemption of the Notes. If the Noteholders receive principal on the Notes earlier than expected, the Noteholders may not be able to reinvest the principal at a similar rate of return.

Alternatively, a decrease in convenience usage may reduce the principal payment rate on the accounts. This could result in the Noteholders receiving the principal on the Notes later than expected.

Prospectus B. Linguistic complexity (Gunning's Fog index): 28.65. Mean sentence length: 42.14.

Under the terms of the Credit Card Agreements, faster than expected rates of principal repayment on the Receivables will, other than during the Revolving Period, cause the Issuer to make payments of principal on the Notes earlier than expected and will shorten the weighted average maturity of the Notes, and slower than expected rates of principal repayment on the Receivables will cause the Issuer to make payments of principal on the Notes later than expected and will lengthen the weighted average maturity of the Notes. Faster or slower repayment of principal on the Receivables may occur as a result of (i) repayments of Receivables by Borrowers at a faster rate than historic payment behaviour; (ii) repayment by the Borrowers above the minimum required repayment; (iii) delinquency and defaults on Receivables and varying speed of recoveries received on defaulted Receivables, and (iv) repurchases by the Originator of any Receivables. A wide variety of economic, social and other factors will influence the rate of repayments on the Receivables. No prediction can be made as to the actual repayment rates that will be experienced on the Receivables (see "***Characteristics of the Receivables - Repayment methods***").

If principal is paid on the Notes of any Class earlier than expected due to faster than expected repayments on the Receivables, Noteholders may not be able to reinvest the principal in a comparable security with an effective interest rate equivalent to the interest rate on the Notes of any Class. Similarly, if principal payments on the Notes of any Class are made later than expected due to slower than expected repayments on the Receivables, Noteholders may lose reinvestment opportunities. Noteholders will bear all reinvestment risk resulting from receiving payments of principal on the Notes of any Note Series earlier or later than expected.

Figure A.5: Example of linguistic complexity of ABS prospectuses

This figure illustrates linguistic complexity in the risk factor section in two different ABS prospectuses. Linguistic complexity is measured as Gunning's Fog index. See Section IV.1 for further details on our measure of linguistic complexity.

Table A.1: Sample selection

	Deal level	Security level
Data available under the LLI	**1,406**	**6,691**
Less		
Prospectus missing	218	1,775
Prospectus not in English	158	505
Issuer not clearly identifiable	16	110
Observations with available prospectuses	**1,014**	**4,301**
Less		
Security-level rating missing	50	1,432
Coupon prices or performance missing	74	363
Control variables missing	24	37
Observations with available prospectuses and security data	**866**	**2,469**

This table reports our sample selection procedure. When excluding observations for which we do not have security-level rating data, it is important to note that the number of deals excluded does not necessarily correspond to the number of securities dropped. This is due to the fact that ratings may be available for some securities of a deal, but not for others.

Table A.2: Variable definitions

Variable	Unit	Description	Data source
Textual disclosure quality and quantity (deal level)			
Boilerplate language	%	Share of words in an ABS prospectus that come from sentences including boilerplate phrases. Boilerplate phrases are defined as tetragrams that appear in at least 33.33% of all prospectuses within the same asset class. For details see Section IV.	EDW, own calculation
Linguistic complexity		Gunning's Fog index. For details see Section IV.	EDW, own calculation
Prospectus length		Natural logarithm of the number of words in an ABS prospectus. For details see Section IV.	EDW, own calculation
Ex post performance (security level)			
Interest shortfall	%	Maximum cumulative difference between the contractually agreed upon interest payment and the actual interest payment for a security since issuance divided by the outstanding security volume.	EDW, S&P Global, own calculation
Investors' pricing (security level)			
Yield spread	%	ABS coupon payments above the reference interest rate at origination if the coupon is floating. In case of a fixed interest, *Yield spread* is calculated by the initially determined interest rate minus the risk-free rate with the most suitable maturity. The risk-free rate is defined by the ECB yield spread index of all sovereign bonds, which are "AAA" rated in the Euro area.	ECB, EDW, FRED, S&P Global, Refinitiv Datastream, own calculation
Controls (deal level)			
Interest rate	%	The mean interest rate of the loans in the respective underlying loan portfolio.	EDW, own calculation
Number of securities		Number of securities in the ABS deal.	EDW, S&P Global, own calculation
Rating disagreement	0/1	Indicator variable equal to one if the ratings of the three CRAs Fitch, Moodys, and S&P differ for at least one security in the ABS deal at the time of issuance.	S&P Global, own calculation

Table A.2: Variables definitions (continued)

Variable	Unit	Description	Data source
Controls (security level)			
Security width	%	Share of the security in the total volume of the ABS pool at the time of issuance.	EDW, S&P Global, own calculation
Security size		Natural logarithm of the principal balance of the security at the time of issuance.	EDW, S&P Global
Security term		Natural logarithm of the time to maturity of the security in years.	EDW, S&P Global
Excess interest	%	Ratio of the maximum amount of excess interest and the principal balance.	S&P Global, own calculation
Subordination	%	Volume, which is subordinated to the respective security, divided by the total volume of the transaction.	EDW, S&P Global, own calculation

This table presents the definitions of the variables used in our analyses. The variables refer to the deal and security level. In the third column, the data sources of each variable are listed.

Table A.3: Correlation matrix

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1)	*Boilerplate language*	1.00												
(2)	*Linguistic complexity*	0.65	1.00											
(3)	*Prospectus length*	0.62	0.63	1.00										
(4)	*Interest shortfall*	-0.10	-0.07	-0.07	1.00									
(5)	*Yield spread*	-0.02	-0.03	0.10	0.17	1.00								
(6)	*Interest rate*	0.34	0.19	0.24	-0.03	0.12	1.00							
(7)	*Number of securities*	-0.01	0.05	0.27	-0.01	-0.00	-0.04	1.00						
(8)	*Rating disagreement*	-0.33	-0.32	-0.25	0.08	0.10	-0.23	-0.04	1.00					
(9)	*Security width*	0.09	0.06	-0.02	-0.07	-0.39	0.05	-0.21	-0.20	1.00				
(10)	*Security size*	0.07	0.04	-0.00	-0.08	-0.46	0.02	0.08	0.21	0.74	1.00			
(11)	*Security term*	-0.32	-0.16	-0.14	0.04	0.01	-0.39	0.15	0.24	-0.15	0.07	1.00		
(12)	*Excess interest*	0.06	0.18	0.20	-0.03	0.37	0.25	0.02	-0.04	-0.27	-0.37	-0.136	1.00	
(13)	*Subordination*	0.13	0.10	0.14	-0.07	-0.21	0.01	0.19	-0.08	0.05	0.34	0.20	-0.13	1.00

This table reports pairwise Pearson correlation coefficients for the variables used in our empirical analysis. Variables are defined in Table A.2.

Table A.4: Examples for boilerplate language in ABS prospectuses

Asset class	Phrase
Auto	cause [to be] communicated [an] invitation [or] inducement
	independently assess [and] determine [the] sufficiency
	certain jurisdictions may [be] restricted
	received [by it in] connection [with the] issue [or] sale
	may [be] offered [,] sold [or] delivered
Consumer	material net economic interest
	principal amount outstanding [of the] notes
	authorised [to] give [any] information [or to] make
	engage [in] investment activity within
	immediately preceding payment date
Credit card	ability [of the] issuer [to] meet [its] obligations
	ability [to] make payments [on the] notes
	able [to] ascertain [from] information published
	assess [and] determine [the] sufficiency [of the] information
	act [or the] applicable laws [of other] jurisdictions
Leasing	limited recourse obligations [of the] issuer
	nature imposed [,] levied [,] collected
	care [to] ensure [that such is the] case [, such] information
	case [only if and to the] extent [that] payments [of a] higher
	certain jurisdictions may [be] restricted
Residential mortgage	engage [in] investment activity within
	taken reasonable care [to] ensure
	respect [to] anything done [by it in] relation
	give [any] information [or to] make [any] representation
	materially prejudicial [to the] interests [of the] noteholders
SME	circumstances [that will] result [in] compliance [with any] applicable
	decisions given [,] expressed [,] made
	likely [to] affect [the] import [of such] information
	omit anything likely [to] affect
	shall [be] binding upon [all the] noteholders

This table reports five boilerplate phrases in ABS prospectuses per asset class in our sample. Common words excluded during the preprocesing of the prospectuses are in square brackets.

Table A.5: Textual disclosure quality and quantity and the yield spread: Alternative measures

	Boilerplate language	Linguistic complexity	Prospectus length
	(1)	(2)	(3)
Bank size	-0.00261	0.0153	0.0159*
	(0.0035)	(0.0437)	(0.0082)
Liquidity	0.000565	-0.000926	0.00288**
	(0.0004)	(0.0038)	(0.0011)
ROE	0.000359	0.0110**	0.000857
	(0.0004)	(0.0050)	(0.0013)
Tier-1 capital ratio	-0.00220**	-0.0201*	-0.00126
	(0.0009)	(0.0104)	(0.0034)
NPL ratio	0.00199*	0.0120	0.00147
	(0.0010)	(0.0090)	(0.0034)
Lawyer FE	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
N	385	385	385
Adj. R^2	0.6745	0.6212	0.5569

The table displays the results of the analysis to which extend the measures of disclosure quality and quantity in ABS prospectuses are related to the characteristics of the issuing bank. *Bank size* is defined as the logarithm of banks' total assets, *Liquidity* by the ratio of liquid assets over total assets, *ROE* represents the net income divided by banks' equity, *Tier-1 capital ratio* describes the regulatory equity ratio, and *NPL ratio* measures the ratio of non-performing over total loans. The analysis is conducted on the deal level, in line with the analysis in Section V.7. The table reports standard errors clustered with respect to the bank level in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A.6: Textual disclosure quality and quantity and the yield spread: Alternative measures

	Yield spread			
	(1)	(2)	(3)	(4)
Boilerplate language	-1.858***			-1.779***
	(0.678)			(0.669)
Linguistic complexity		-0.041		-0.041
		(0.026)		(0.025)
Prospectus length			0.356***	0.386**
			(0.118)	(0.117)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,469	2,469	2,469	2,469
Adj. R^2	0.60	0.60	0.60	0.61

This table displays the results of the analysis how alternative measures of textual disclosure quality and quantity in ABS prospectuses are related to the initial yield spread demanded by investors (*Yield spread*). The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table A.7: Textual disclosure quality and quantity and the yield spread: Floating coupons

	Yield spread			
	(1)	(2)	(3)	(4)
Boilerplate language	-0.681**			-0.642**
	(0.277)			(0.276)
Linguistic complexity		-0.011		-0.013
		(0.020)		(0.021)
Prospectus length			0.445***	0.446***
			(0.108)	(0.108)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Observations	1,838	1,838	1,838	1,838
Adj. R^2	0.65	0.64	0.65	0.65

This table displays the results of the analysis how measures of textual disclosure quality and quantity in ABS prospectuses are related to the initial yield spread demanded by investors (*Yield spread*), when only including ABS with floating coupons. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

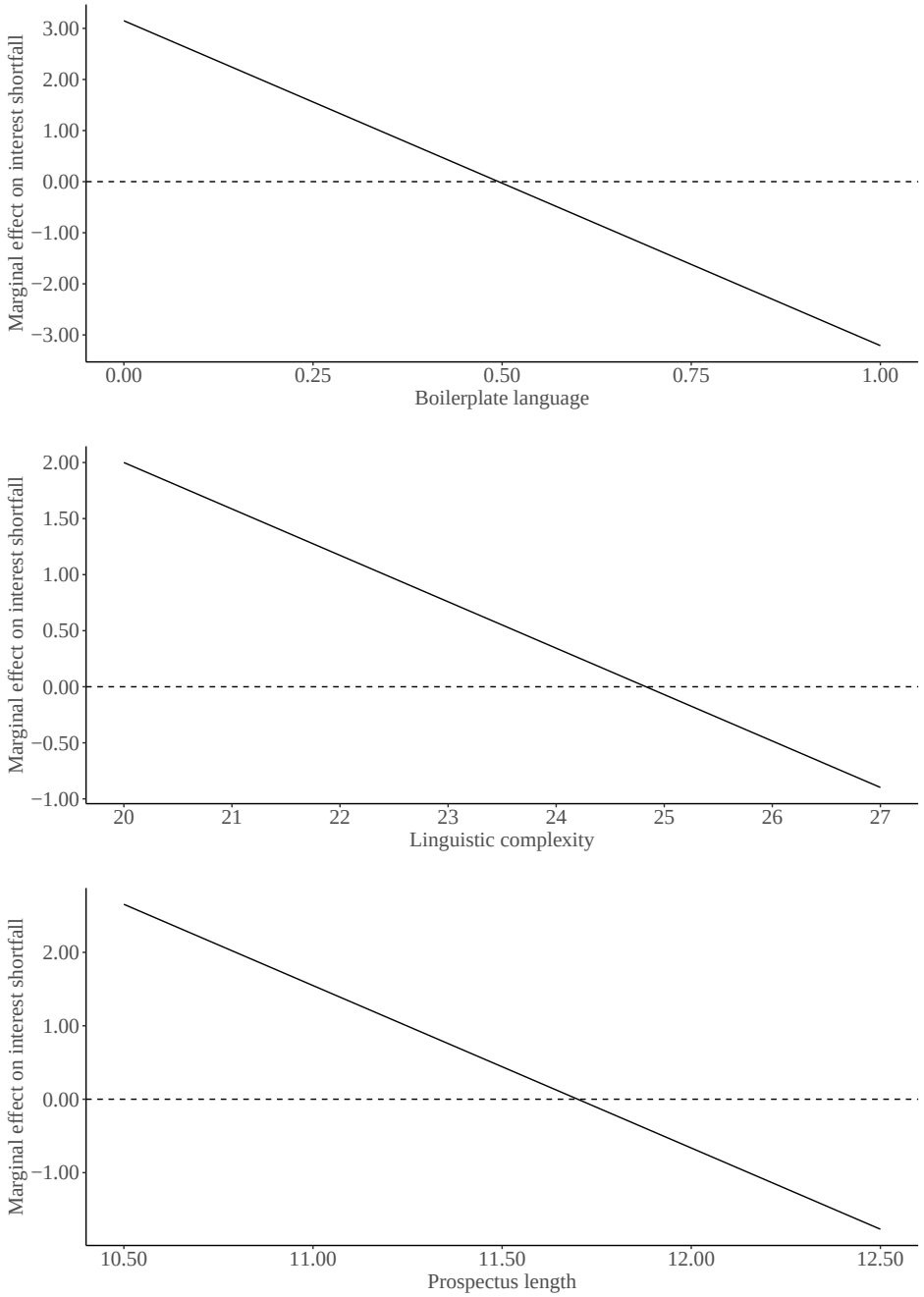

Figure A.6: Marginal effect of the yield spread on the interest shortfall

This figure displays the marginal effect of the initial yield spread demanded by investors (*Yield spread*) on the ex post performance of ABS (*Interest shortfall*), depending on the share of boilerplate language, the linguistic complexity, and the prospectus length as measures of textual disclosure quality and quantity.

Table A.8: Textual disclosure quality and quantity and the predictive ability of the yield spread: Alternative measures

	Interest shortfall			
	(1)	(2)	(3)	(4)
Yield spread	1.289***	8.574***	13.953***	12.416***
	(0.427)	(3.245)	(3.768)	(3.593)
Boilerplate language x Yield spread	-10.220***			-5.061*
	(3.210)			(2.927)
Linguistic complexity x Yield spread		-0.422**		0.0126
		(0.167)		(0.231)
Prospectus length x Yield spread			-2.513***	-2.199***
			(0.672)	(0.756)
Boilerplate language	11.760***			8.895**
	(3.957)			(3.707)
Linguistic complexity		0.496***		0.070
		(0.143)		(0.119)
Prospectus length			2.381***	2.041***
			(0.666)	(0.692)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Coupon type FE	Yes	Yes	Yes	Yes
Observations	2,469	2,469	2,469	2,469
Adj. R^2	0.20	0.20	0.21	0.22

This table displays the results of the analysis how the interaction between alternative measures of textual disclosure quality and quantity on the one hand and the initial yield spread demanded by investors (*Yield spread*) on the other hand is related to the ex post performance of ABS (*Interest shortfall*), indicating the predictive ability of yield spreads at security issuance. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table A.9: Textual disclosure quality and quantity and the predictive ability of the yield spread: Floating coupons

	Interest shortfall			
	(1)	(2)	(3)	(4)
Yield spread	4.571***	36.760***	13.880***	3.031
	(1.204)	(10.672)	(5.235)	(7.221)
Boilerplate language x Yield spread	-8.305***			-9.221***
	(2.147)			(2.845)
Linguistic complexity x Yield spread		-0.538**		0.400
		(0.218)		(0.348)
Prospectus length x Yield spread			-3.108***	-0.654
			(0.907)	(0.853)
Boilerplate language	9.016***			10.010***
	(2.515)			(3.079)
Linguistic complexity		0.620***		-0.246
		(0.188)		(0.198)
Prospectus length			3.817***	1.383**
			(1.073)	(0.575)
Controls	Yes	Yes	Yes	Yes
Deal orig. year FE	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Rating FE	Yes	Yes	Yes	Yes
Observations	1,838	1,838	1,838	1,838
Adj. R^2	0.31	0.26	0.28	0.31

This table displays the results of the analysis how the interaction between measures of textual disclosure quality and quantity in ABS prospectuses on the one hand and the initial yield spread demanded by investors (*Yield spread*) on the other hand is related to the ex post performance of ABS (*Interest shortfall*), indicating the predictive ability of yield spreads at security issuance, when only including ABS with floating coupons. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table A.10: Textual disclosure quality and quantity and regulation: Longer time period

	Entire prospectus			Risk factor section		
	Boilerplate language	*Linguistic complexity*	*Prospectus length*	*Boilerplate language*	*Linguistic complexity*	*Prospectus length*
	(1)	(2)	(3)	(4)	(5)	(6)
Boilerplate language (high) x *Post*	-0.041**			-0.049***		
	(0.019)			(0.015)		
Linguistic complexity (high) x *Post*		-0.295			-0.967***	
		(0.235)			(0.321)	
Prospectus length (high) x *Post*			-0.123**			-0.390***
			(0.054)			(0.107)
Boilerplate language (high)	0.081***			0.037*		
	(0.016)			(0.019)		
Linguistic complexity (high)		0.976***			1.551***	
		(0.304)			(0.395)	
Prospectus length (high)			0.226***			0.364***
			(0.069)			(0.082)
Deal orig. year FE	Yes	Yes	Yes	Yes	Yes	Yes
Asset class FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	601	649	575	788	726	592
Adj. R^2	0.88	0.75	0.76	0.77	0.67	0.91

This table displays the results of the analysis how EU regulations in 2017 affected measures of textual disclosure quality and quantity in ABS prospectuses in general (columns (1) to (3)) and in the risk factor section (columns (4) to (6)) using a longer sample. The table reports standard errors clustered at the ABS deal level in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

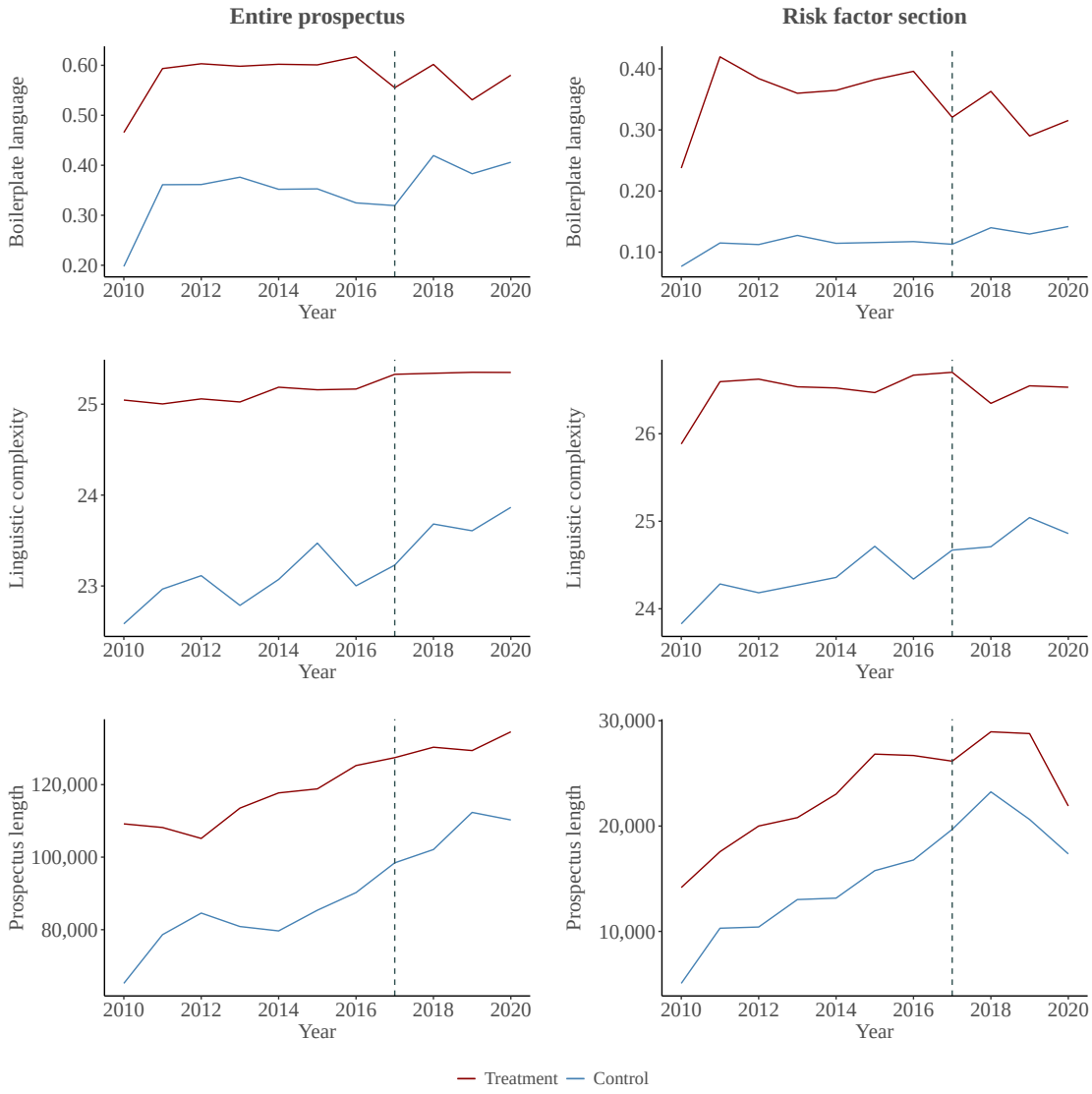

Figure A.7: Textual disclosure quality and quantity and regulation

This figure displays the means of the textual disclosure quality and quantity measures based on the entire prospectus (left part) and based on the risk factor section (right part) for the treatment and control groups over time. The date of the considered EU regulations, 2017, is marked with vertical lines.