BANK OF ENGLAND

Staff Working Paper No. 834

The language of rules: textual complexity in banking reforms

Zahid Amadxarif, James Brookes, Nicola Garbarino, Rajan Patel and Eryk Walczak

October 2019

BANK OF ENGLAND

Staff Working Paper No. 834
The language of rules: textual complexity in banking reforms

Zahid Amadxarif,[1] James Brookes,[2] Nicola Garbarino,[3] Rajan Patel[4] and Eryk Walczak[5]

Abstract

The banking reforms that followed the financial crisis of 2007–08 led to an increase in UK banking regulation from almost 400,000 to over 720,000 words, and to concerns about their complexity. We define complexity in terms of the difficulty of processing linguistic units, both in isolation and within a broader context, and use natural language processing and network analysis to calculate complexity measures on a novel dataset that covers the near universe of prudential regulation for banks in the United Kingdom before (2007) and after (2017) the reforms. Linguistic, ie textual and network, complexity in banking regulation is concentrated in a relatively small number of provisions, and the post-crisis reforms have accentuated this feature. In particular, the comprehension of provisions within a tightly connected 'core' requires following long chains of cross-references.

Key words: Complexity, CRR, natural language processing, networks, regulation, supervision.

JEL classification: D85, E58, G18, G28, K23.

(1) Bank of England. Email: zahid.amadxarif@bankofengland.co.uk
(2) Bank of England. Email: james.brookes@bankofengland.co.uk
(3) Bank of England. Email: nicola.garbarino@bankofengland.co.uk
(4) Bank of England. Email: rajan.patel@bankofengland.co.uk
(5) Bank of England and University College London. Email: eryk.walczak@bankofengland.co.uk

The views in this paper are the authors' and do not necessarily reflect those of the Bank of England or its policy committees. The authors are extremely grateful to David Bholat, Julia Black, Jean-Edouard Colliard, Jon Danielsson, Stephen Dickinson, Bill Francis, Paul Grout, Andy Haldane, Martin Lodge, Mirko Moscatelli, Victoria Saporta, Misa Tanaka and seminar participants at the London School of Economics, the European Association of Law and Economics, the Federal Reserve Board, the ESCB Research Cluster 3 conference at Banco de España, the OECD and the Bank of England for helpful comments. We also thank Joel Mundy and Matthew Naylor for their excellent research assistance. Any remaining errors or omissions are the authors' own.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Email publications@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

Regulatory reforms are often a reaction to disasters (air crashes, gas explosions, oil spills, financial crises) that expose the complexity of social systems. Inquiries often reveal that, although the immediate cause is often an operating failure (Perrow, 2011), disasters are ultimately linked to failures in regulatory design (Black, 2014). When trying to control social systems, rule drafters make choices about how to encode real world complexity in legal texts. These choices affect the cognitive processing required to enforce and comply with rules—and hence matter for the rules' effectiveness in reducing the likelihood of disasters. In this paper, we interpret regulatory complexity in terms of *processing complexity*, by using techniques from natural language processing (NLP) and network analysis and applying them to a major regulatory reform of the past decade—the new international banking rules that followed the Great Financial Crisis.

The wave of banking reforms resulted in a substantial revision of regulatory texts. Growth in the length of the post-crisis prudential regulatory framework and the potential interactions between different rules have led to concerns that the commitment to prevent another crisis may have led to excessive regulatory complexity (Romano, 2012; Haldane and Madouros, 2012; Greenwood et al., 2017).[1]

Regulators do not aim to produce complex rules; regulatory complexity is usually the by-product of their attempts to manage complex systems, such as financial markets.[2] Tightening rules to reduce risks, such as the risk of a financial crisis, requires taking into account more factors and eventualities, increasing complexity (Ehrlich and Posner, 1974; Battigalli and Maggi, 2002). Similar to environmental pollution, this view frames complexity as an external cost that is incurred in the production of an activity—in this case, regulation—that has social benefits. Complex regulation will require more effort and information to comprehend, and it can create uncertainty if full comprehension is not achieved, because agents find it inefficient and settle for partial comprehension, preferring to use simple heuristics (Gigerenzer and Selten, 2002). Sometimes uncertainty cannot be fully eliminated because rules are by nature generalisations, which have to contain a degree of vagueness to adapt to new circumstances (Black, 1997). In this case, complexity may be seen as the price to pay to leave room for regulatory discretion that can draw on a wider information set (e.g. information that becomes available only after the rule is drafted, see Endicott (2005)). Asymmetric information resulting from differences in agents' comprehension abilities can result in lower accountability or favour regulatory capture by the industry (Glaeser and Shleifer, 2002; Hakenes and Schnabel, 2014).

Like pollution, regulatory complexity creates externalities, in the sense that the drafter bears only part of the total costs, while agents that have to comply with the rules (banks) or enforce

[1] For useful reviews of banking reforms since the crisis see Duffie (2017), Claessens et al. (2017), Aikman et al. (2018), and Tarullo (2019).

[2] A different view is that complexity is the result of deliberate lobbying by large incumbents to increase barriers to entry (see for example, Posner (1975)). Gai et al. (2019) provide an overview of explanations for regulatory complexity in financial markets.

them (supervisors) bear the rest. In order for the regulator to take into account the full social costs of complexity, these costs must be measured, so that they can be included in the overall calculation of social costs and benefits. The nature of regulatory frameworks as complex systems—ones made up of large numbers of parts that interact in non-simple ways with unexpected results (Simon, 1991)—makes this particularly challenging. Interconnections between individual provisions make it harder for regulated parties to understand policy intent, and for regulators to understand the full costs and benefits of using or changing individual rules (Ruhl and Salzman, 2002). Recently, network analysis has started to map interconnections in legal contexts—e.g. US Supreme Court opinions (Fowler et al., 2007) and the EU's Official Journal (Koniaris et al., 2017)—but we are not aware of specific applications to prudential regulation for banks. Instead, studies of the impact of the post-crisis reforms typically focus on individual policies, abstracting from the rest of the regulatory framework.[3]

Regulatory complexity can be defined in different ways: for example, as the number of facts that are taken into account (Kaplow, 1995), as the number of parameters a regulation might require a bank to estimate (Aikman et al., 2014), or as the resulting balance sheet complexity (Goldberg and Meehl, 2019). In this paper, we define complexity in terms of the processing difficulty encountered when comprehending a particular linguistic unit (Gibson, 1998), e.g. a single regulatory provision. Dimensions of processing difficulty for a provision include its length, lexical diversity, use of conditional statements, and the overall readability of its sentences (we define this as "local" complexity). Some processing difficulties can only be resolved after accessing information outside the immediate context of the provision—for instance, cross-references or regulatory precedents needed to understand a provision's intent ("global" complexity). We use natural language processing and network analysis techniques to measure these dimensions of local and global complexity, and apply these measures to a novel dataset we construct ourselves.

To anticipate our main result, we find that linguistic complexity in banking regulation is concentrated in a relatively small number of provisions. Starting from the simplest provisions, our measures of complexity increase slowly, but then pick up rapidly as we approach the last ten per cent of most complex provisions. This stylised fact has been accentuated by the post-crisis reforms, which have resulted in the rise of highly complex provisions, in particular a tightly connected core. Our results could reflect increasing marginal linguistic complexity: the linguistic complexity of individual provisions increases as rule drafters seek to encode additional real world complexity.

In the remainder of this introduction, we outline our dataset, measure and key results. The first step of our study is to create a dataset that i) captures legal sources comprehensively; ii) allows like-for-like comparison between the post- and pre-crisis frameworks; and iii) captures the entire structure of cross-references within the regulatory framework (to facilitate network analysis). We create a new dataset that includes the near universe of prudential legal obligations and supervisory guidance that applied to UK banks in 2007 and 2017.

[3]The literature evaluating banking reforms is vast. Recent examples include Gropp et al. (2018), Behn et al. (2016), Juelsrud and Wold (2018), and Galati and Moessner (2018).

Our dataset captures changes in both the scope of what regulators seek to control and in the legal architecture. Since 2007, efforts to address shortcomings in pre-crisis banking regulation have introduced some new regulatory concepts—e.g. UK rules on individual accountability for senior managers within banks—and significantly revised existing concepts, e.g. tightening the definition of eligible capital instruments. The balance between EU and UK legal obligations changed with the introduction of the Capital Requirements Regulation (CRR) and accompanying European Banking Authority (EBA) Technical Standards, which are directly binding on banks and do not require implementation through national legislation. Within the UK, a new Prudential Regulation Authority took responsibility for prudential regulation in 2013.

The second step of our study is to construct measures that capture the change in the language and structure of regulatory texts. Measures of local complexity focus on difficulties in processing the language of individual provisions. *Length* focuses on the amount of linguistic material a reader has to retrieve, integrate and keep in working memory for comprehension. *Lexical Diversity* captures the proportion of new concepts (comprehension is facilitated when concepts are repeated). *Conditional* clauses and expressions require integrating different exceptions or constructing hypothetical events.

Global comprehension happens after the reader has partially processed the text contained in a provision, and has to access additional contextual information to complete full, successful comprehension. We use two measures, derived from network analysis, that capture interconnections between provisions via cross-references. *Degree* counts the number of direct connections (cross-references) from and to a provision. *PageRank* summarises the centrality of a provision within the regulatory framework and captures the effects of chains of indirect connections via cross-references. The "ripple effects" from regulatory changes—and the non-linearity of policy-transmission mechanisms—are likely to be greatest for provisions with the highest PageRank scores.

In addition, vague terms such as "reasonable" or "adequate" may become clear only in the light of precedents or common practice in a specific market. We capture the extremes of a hypothetical *vagueness-precision* spectrum by exploiting i) a hand-crafted lexicon of vague terms used in prudential regulation, and ii) precise numerical indicators of currency and per cent values.

Local and global complexity could be defined at different levels of aggregation, i.e. the linguistic unit of reference. Legal texts are structured hierarchically in increasingly larger units such as articles, chapters, parts etc. We present results for two units of observations: the lowest level of aggregation for each source document, which we label as individual "provisions", and a higher level of aggregation ("chapters") that captures wider regulatory topics. Individual provisions appear a more natural level to separate the local and global complexity, but we also show chapter-level results to clarify how aggregation can change the results of the analysis.

We start by discussing results for provision-level measures. Distributions for both local and global measures indicate that linguistic processing difficulties are not spread evenly across

provisions. Most provisions are relatively simple, and complexity is concentrated at the top end of the distributions. The post-crisis reforms accentuated this feature, introducing additional complexity at the top end of the distribution, and creating highly complex provisions in the top deciles.

The rise of highly complex provisions is most visible for *global* measures based on cross-references. Post-reform, a core of tightly connected provisions emerges. The comprehension of provisions in this core requires accessing and organising information that is articulated over long chains of cross-references. At one end of the longest chains we find key provisions (e.g. on the definition of default and minimum capital requirements) in CRR, which was the main instrument through which international banking reforms were enacted in the EU.

The comprehension of post-reform provisions seems to require also accessing more contextual information that is external to the regulatory framework. Relatively fewer rules include specific numerical values in 2017 (indicators of precise regulatory thresholds), and vague terms appear relatively more frequently in 2017 guidance.

Highly complex provisions appear also in the distributions for *local* measures for binding rules. Compared to 2007, the top deciles of rule provisions in 2017 contain longer provisions, with higher lexical diversity, and conditionality. This indicates that post-crisis reforms have generated a (small) set of rules that require retrieving a lot of material, containing several concepts and exceptions. Guidance, on the opposite, has changed less as a result of the reforms.

Results at chapter level broadly confirm those at provision level. The most notable difference is in the global measure based on cross-references.[4] The PageRank distribution for 2017 is above the distribution for 2007 only in top decile. CRR appears to account for almost all the interconnections in the 2017 chapter-level network.

Our results contribute directly to the literature that evaluates the complexity of bank regulation, by integrating approaches from network analysis and NLP. Studies of network complexity in financial markets have focused on the contagion effects that derive from networks of interbank exposures (Battiston et al., 2012, 2016; Bardoscia et al., 2017). Enriques et al. (2019) have argued that bank regulation should be "network-sensitive" and take into account these interconnections. A different literature, rooted in evolutionary economics, argues that regulatory complexity can lead to "kludging": incremental policy changes leading to a sub-optimal web of rules that is difficult to reform (Ely, 2011; Kawai et al., 2018). Testing these theoretical claims requires measures that capture properties of banking rules as networks, similarly to studies on legal networks in non-financial areas, e.g. U.S. Supreme Court opinions (Fowler et al., 2007). Our results show that network complexity has increased post-crisis, and is concentrated in a tightly connected core.

Research in finance and economics that makes use of the notion of textual complexity is rapidly

[4]Note that the network is more interconnected at chapter than at provision level. By aggregating individual provisions, chapters effectively create links between the provisions that they contain.

growing, but complexity is seldom adequately defined.[5] For instance, for Bushee et al. (2018), textual complexity involves an increase in "information processing costs", but nowhere do the authors state what exactly these processing costs are.[6] We explain how measures of local and global complexity are grounded in linguistic research, and how they can be interpreted in the light of legal and economic literature. We show that the increase in local complexity measures is concentrated at the top end of the distributions, while complexity for the remaining provisions has remained fairly constant over time, suggesting that rule-writers are aware of the linguistic burden and try to limit it for individual rules.

Two studies in particular have helped shaped our analysis of regulatory complexity. Our measures of network complexity are inspired by Li et al. (2015)'s study of the US Code of Law, and our distinction between local and global complexity is similar to Colliard and Georg (2018)'s idea that regulators can use "built-in functions" to manage complexity.[7] We combine these two approaches, and motivate our measures in linguistic terms. More importantly, our dataset does not rely on a single legal source, but captures a broad enough set of sources to facilitate comparison of the prudential regulatory framework for banks before and after the financial crisis and the following reforms. To our knowledge, this is the first study to consolidate these sources. We plan to release the code we used to construct our dataset.[8]

Our focus on the language of legal texts is also informed by the legal debate on precise rules versus vague standards (for an overview see Cooter and Ulen (2016)). Looking ahead, this long-standing debate has implications for understanding how artificial intelligence (AI) can affect law and regulation (Kroll et al., 2016; Levmore and Fagan, 2019), which in turn can have implications for banks' risk management and financial stability (Danielsson et al., 2017). AI is expected to perform well when regulation is expressed precisely, but it is less clear how effective it can be at interpreting vague standards. Our local/global distinction uses linguistic concepts to operationalise different dimensions of complexity. We interpret vagueness as a form of global processing complexity, whose (partial) resolution depends on the readers' ability to access available information, e.g. from precedents and market knowledge.

The remainder of the paper is structured as follows: Section 2 discusses our legal sources and the resulting dataset; Section 3 describes our complexity measures; Section 4 presents our results; and Section 5 concludes.

[5] Gentzkow et al. (2017) provides an overview of the uses of textual analysis in economics. NLP techniques have been used to analyse legal texts (Katz and Bommarito, 2014; Ash, 2015; Li et al., 2015), central bank communication (Born et al., 2014; Bholat et al., 2015; Chakraborty and Joseph, 2017; Haldane, 2017) and corporate reporting (Loughran and McDonald, 2014; Leuz and Wysocki, 2016; Loughran and McDonald, 2016; Hoberg et al., 2014; Chen et al., 2014; Engelberg et al., 2012; Garcia, 2013; Da et al., 2014; Buehlmaier and Whited, 2018; Handley and Li, 2018).

[6] Similar criticisms can be leveled at Brochet et al. (2012); Guay et al. (2016); de Souza et al. (2019).

[7] Colliard and Georg (2018) identify specialised vocabulary as built-in functions summarising legal or market knowledge.

[8] We have released key elements of our code alongside this paper and will release more at a later stage.

2 Regulatory texts and data

2.1 Elements of our dataset

We create a new dataset that includes the near universe of prudential legal obligations, and supervisory guidance, that apply directly to UK-authorised banks.

For the purposes of this paper, we describe this as the "prudential regulatory framework" and define key terms as follows:

- *Prudential*: regulation relevant for the Prudential Regulation Authority's (PRA) scope of powers. These are largely financial requirements concerning the valuation of assets and liabilities and the amounts of capital and liquidity to be held and their quality; together with associated non-financial requirements concerning risk management, governance and reporting.
- *Legal obligation*: any mandatory requirement that binds directly on a UK-authorised firm, such as PRA rules.
- *Guidance*: regulatory authorities' expectations for how firms should comply with their legal obligations or should expect a given process to be dealt with.
- *Banks*: a firm subject to the EU's Capital Requirements Directive, i.e. deposit-takers including building societies (but not credit unions).

We seek to construct a dataset that: 1) is reasonably comprehensive with respect to post-crisis prudential regulation; 2) captures equivalent scopes of prudential regulation from before and after the crisis to identify post-crisis changes robustly; and 3) facilitates analysis of prudential regulatory frameworks as networks.

We use 16 November 2007 and 16 November 2017 as observation dates for the pre- and post-crisis prudential regulatory frameworks. These allow us to capture the cumulative effect of the first wave of post-crisis banking reforms.

The rest of this section describes: key post-crisis changes in the prudential regulatory framework which influence how we built the dataset and interpret our measures; elements of the dataset; its use for network analysis; and our choices about units of observation.

2.2 Prudential regulation for banks in the EU and the UK: relevant post-crisis changes

Three key dimensions of change in the prudential regulatory framework were:

- Change (expansion) in the scope of what regulators seek to control;
- Change in the balance between EU and UK legal obligations; and
- Responsibility for prudential regulation in the UK moving from the Financial Services Authority (FSA) to the PRA (a subsidiary of the Bank of England).

We discuss each change in turn.

Change in the scope of what regulators seek to control

Since 2007, efforts to address shortcomings in pre-crisis regulation have introduced several new concepts—and revised existing concepts—in the prudential regulatory framework.

Some originate from changes to international standards. In particular, the Basel III reform package: introduced a leverage ratio requirement, two new requirements for mitigating excessive liquidity risk and maturity transformation (Liquidity Coverage Ratio and Net Stable Funding Ratio), and macroprudential tools; and it revised existing provisions on the definition of capital and the calculation of capital requirements. And the Financial Stability Board set out key attributes of effective resolution regimes for financial institutions.

Other concepts reflect UK own-initiative policy changes. In particular, policy on: *ring-fencing* (the structural separation of banks' retail- and wholesale / investment-banking activities); *strengthening accountability* (Senior Managers and Certification Regime, corporate governance, and remuneration); and macroprudential tools for the housing market.

Change in the balance between EU and UK legal obligations

The EU's aim to deliver a "unified regulatory framework for the EU financial sector"—a Single Rulebook—has fundamentally changed how prudential rules and guidance are articulated in the UK.

In 2007, the EU used Directives to articulate its expectations for banking provisions. Directives express goals for EU countries to achieve, but leave it up to individual countries to legislate to reach these goals. Moreover, these directives were generally "minimum harmonising", meaning that a jurisdiction like the UK had substantial rule content which was additive to directive transposition. Finally, the UK's 2007 prudential regulatory framework was authored entirely by UK lawmakers; and while a significant proportion of it implemented provisions in EU Directives, a single document—the FSA's Handbook—gave UK banks all the rules and guidance they needed to comply with.

The financial crisis led the EU to reform this framework. Two changes had material effects:

- To consolidate the Single Market and reduce potential for divergence in national rules, the EU used a regulation as well as directives to implement the Basel III reform package agreed by the Basel Committee in 2010. This Capital Requirements Regulation (CRR) is a legislative act which must be applied in its entirety across the EU, and which binds directly on firms without requiring implementing national legislation.
- A new pan-European supervisor (the European Banking Authority, EBA) was also empowered to develop technical standards to specify details which regulations could not cover. Once implemented, a technical standard has the same legal status as a regulation.

So by 2017, the UK's prudential framework had become a patchwork of: 1) EU-authored regulations and technical standards; 2) UK legislation implementing provisions in EU directives; 3) UK-authored provisions covering areas not strictly occupied by the scope of EU rules or the concept of maximum harmonisation; and 4) UK-authored guidance in support of all of the above.

EU regulations are written to harmonise definitions and methodologies across Member States. They emerge from negotiations between Member States and co-decision between the European Parliament, Council and the European Commission; and are drafted by teams of lawyers from different Member States. As the objectives and context for UK own-originated legislation are quite different, we might expect this to produce variation in the properties of language and structure between EU and UK legal texts.

Responsibility for prudential regulation in the UK moving from the FSA to the PRA

The PRA took responsibility for prudential regulation of banks in April 2013. Its statutory objectives differed from the FSA's, and it had a new approach to setting and communicating regulatory policy. Both brought significant change to the prudential regulatory framework.

While the FSA had been responsible for conduct and prudential regulation, the PRA's statutory objective limited its responsibility to prudential regulation. Provisions in the 2007 FSA Handbook included both conduct and prudential standards, but on 1 April 2013, the PRA adopted only the Handbook's prudential aspects.

In the first "approach document" it published as an independent body (April 2013), the PRA said it would aim "to establish and maintain published policy material which is consistent with its objectives, clear in intent, straightforward in its presentation and as concise as possible, so that it is usable by the senior management of firms".

Three changes had material effects:

- Separation of rules and guidance: the FSA had legal power to make rules and guidance,

which it exercised by giving lawyers responsibility for writing guidance as well as rules. It chose to publish its rules and guidance together in its Handbook. The PRA kept rule-making powers only in its Rulebook and chose to publish guidance in the form of standalone, policymaker-authored Supervisory Statements and similar documents.

- Restructuring of the Rulebook: the PRA sought to make prudential rules more tailored and easily searchable for firms by dividing rules into "sectors" (banking and insurance being the key ones) and changing how rules were grouped and labelled. Specifically, the PRA created Rulebook "parts": a new grouping of rules that deal with a similar risk, with part names describing their content (e.g. "Market Risk").

- A new, online Rulebook: the PRA sought to improve the clarity and navigability of PRA rules by launching a new Rulebook website in August 2015.

2.3 Data sources

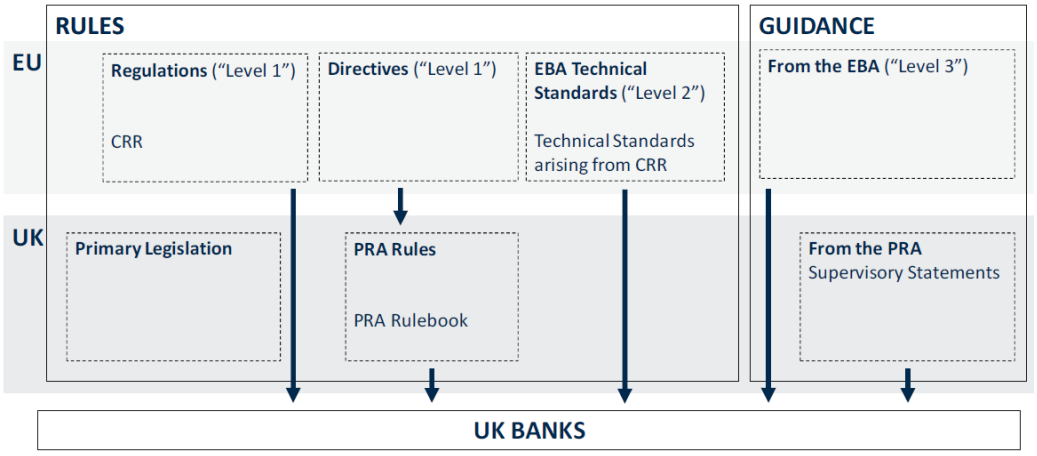
We seek to build a dataset that is reasonably comprehensive for the post-crisis period; and which captures comparable scopes of prudential regulation in the pre- and post-crisis periods to facilitate meaningful comparison.

This required judgments on which sets of source documents to include. And as some documents include provisions that are in and out of scope for our analysis—e.g. prudential (in scope) and conduct (out of scope) provisions, or banking (in scope) and insurance (out of scope) provisions—it required judgments on which provisions within these source documents are relevant for our analysis.

2.3.1 Choice of source documents for the dataset

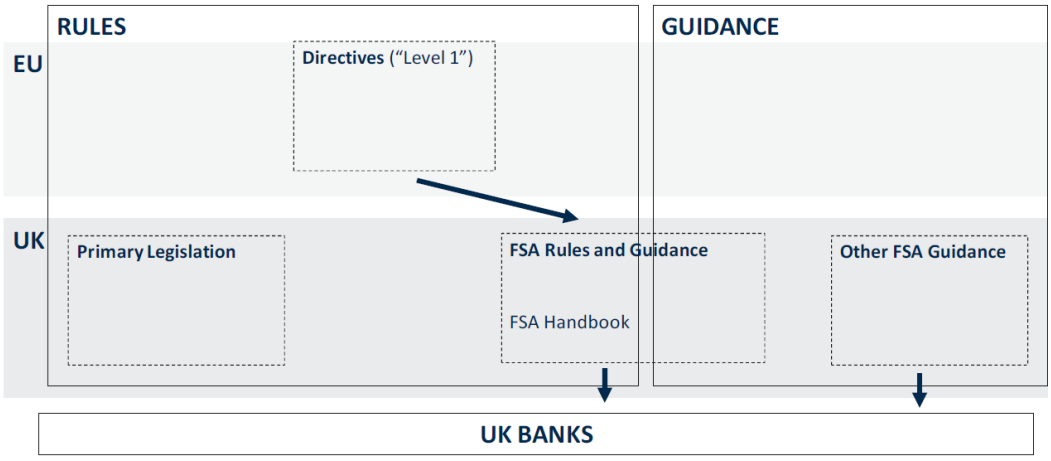
Figures 1 and 2 show the subset of texts that forms the scope of our analysis. The arrows show provisions that apply directly to UK banks. For guidance, these include texts issued by EU and UK supervisory authorities. For rules, these are provisions which are legally binding on UK banks—we restrict our analysis to these texts as they include the most salient prudential rules that have force of law. A consequence is that we exclude the text of EU Directives, e.g. the Capital Requirements Directive which formed part of the EU's legislative package (CRDIV) for implementing the first phase of Basel III reforms. But by including the UK rules which implement them, we capture the text of the legally binding rules they create for UK banks.

Figure 1: **2017 legal sources**



Note:
Arrows show provisions which apply directly to UK banks (EU regulations; EBA Technical Standards; the PRA Rulebook; and supervisory guidance). We capture: 1) the EU Capital Requirements Regulation (CRR); 2) EBA Technical Standards arising from CRR; 3) the PRA Rulebook; and 4) PRA Supervisory Statements.

Figure 2: **2007 legal sources**



Note:
Arrows show provisions which apply directly to UK banks (the FSA Handbook; and other FSA Guidance). We capture the FSA Handbook.

For 2017, at the EU level, we include the EU's Capital Requirements Regulation (CRR) and associated Technical Standards made by the European Banking Authority (EBA) to implement the first phase of Basel III reforms in Europe. At the UK level, we include the PRA Rulebook to capture UK rules implementing relevant EU Directives and UK own-initiative policies. And we include PRA Supervisory Statements, which state expectations for how banks should comply with PRA rules (the key form of supervisory guidance in the UK).

For 2007, we include the FSA's Handbook. It captures the legally binding rules which implemented all relevant EU Directives in the UK; other relevant UK rules; and key supervisory guidance.

Table 1 summarises these judgments. Our choices reflect a trade-off between a desire for comprehensiveness and the analytical cost of turning disparate texts into formats amenable for textual analysis. Future work could expand the scope of our dataset to include other relevant texts, e.g. the EBA's supervisory guidance; and other forms of PRA guidance (e.g. Approach Documents).

Table 1: **Legal sources included in the dataset**
The FSA Handbook was scraped as of 16-11-2007. The PRA Rulebook was scraped as of 16-11-2017.

Included		**Not Included**	
2007	**2017**	**2007**	**2017**
FSA Handbook (banking prudential provisions)	PRA Rulebook (provisions applying to deposit-takers subject to the Capital Requirements Regulation (CRR))	Any provisions applying exclusively to non-banks (e.g. insurers and credit unions). Any conduct-related provisions (in FSA Handbook in 2007 or FCA Handbook in 2017).	
	PRA Supervisory Statements (those applying to CRR firms)	UK primary and secondary legislation other than PRA rules (e.g. Financial Services and Markets Act; Banking Act)	
	Capital Requirements Regulation (CRR, EU 575/2013)	EU directives (e.g. Banking Consolidation Directive and Capital Requirements Directive IV (CRDIV)), as relevant provisions are implemented mainly via the PRA Rulebook	
	EBA Technical Standards arising from CRR		EBA Questions and Answers (Q&A), as we focus on UK-authored guidance (PRA Supervisory Statements) at this stage of the analysis
		Glossary definitions, as at this stage of the analysis, we restrict our analysis to the language and network structure of provisions which articulate policy intent	

2.3.2 Choice of relevant provisions within selected source documents

Within these documents, we restrict our analysis to provisions on prudential requirements for banks ("relevant" provisions) and use web-scraping techniques to consolidate provisions published across multiple online sources into a single dataset.[9]

For 2017, all provisions in the CRR relate to prudential requirements on banks, so we include the entire text. For other texts, we exploit the online interfaces published by the EBA and the PRA to identify relevant provisions:

- EBA Technical Standards: the EBA's Single Rulebook connects individual Technical Standards with the CRR articles mandating them. We used these links to identify the subset of Technical Standards associated with the CRR and included their entire texts.
- PRA rules: the PRA's online Rulebook groups rules according to the "sectors" of firms they apply to. We used this grouping to identify the subset of PRA rules associated with "CRR Firms" (banks) and included their entire texts.
- PRA guidance: the PRA's online Rulebook publishes "Related Links" for each set of rules on a single topic (a Rulebook "Part"). We used these groupings to identify the subset of PRA Supervisory Statements associated with "CRR firms" (banks) and included their entire texts.

For 2007, identifying the comparable set of provisions was more challenging as the FSA's Handbook did not clearly distinguish between: 1) prudential- and conduct-related provisions; and 2) provisions applying to banks versus other sectors of firms.

To identify prudential provisions, we exploit the *FCA and PRA Handbook Designation Instrument 2013* which the FCA and the PRA published as they prepared to take separate responsibility for prudential and conduct regulation in April 2013. It documented their decisions about whether each provision in the FSA Handbook as at 27 February 2013 was prudential, conduct, or both. We also used text-matching techniques to identify provisions that appeared in the FSA Handbook at both 27 February 2013 and at 16 November 2007 (our observation date for 2007 texts).

- For provisions that appeared in the FSA Handbook at both dates, we use the assignment in the Handbook Designation Instrument to identify the prudential subset and included their entire texts.
- For provisions that appeared in the FSA Handbook at 16 November 2017 but were not assigned in the Handbook Designation Instrument (e.g. because they were deleted between 2007 and 2013), we use expert judgment to identify the prudential subset. Table 2 summarises these judgments.

[9] An R package to scrape the PRA Rulebook was developed in the process and is publicly available at https://github.com/bank-of-england/PRArulebook.

Having identified the set of prudential provisions, we use expert judgment to identify the subset relevant for banks. Table 3 shows provisions we judge to be exclusively relevant for non-banks, and hence exclude.

The resulting sets of provisions form the "near universe" of prudential legal obligations, and supervisory guidance from UK authorities, that applied directly to UK-authorised banks in 2007 and 2017.

Table 2: **Judgments made to designate FSA Handbook (2007) provisions as prudential or conduct**

Set of Handbook provisions	**Judged relevant for prudential regulation?**
Credit Unions (CRED)	Yes
Client Assets Sourcebook (CASS)	No
Complaints Against the FSA (COAF)	No
Decision Procedure and Penalties Manual (DEPP)	No
Dispute Resolution: Complaints (DISP)	No
Electronic Money (ELM)	No
ICOB	No
Recognised Investment Exchanges and Recognised Clearing Houses (REC)	No

Note: The sets of provisions in this table are: 1) not captured by the FCA and PRA Handbook Designation Instrument 2013 which designated the majority of FSA Handbook (2007) provisions as either "prudential" or "conduct"; and 2) material in terms of length, so we used expert judgment to designate them.

The sets of provisions judged not relevant for prudential regulation were excluded from our 2007 dataset. We assume that any residual undesignated provisions in the FSA Handbook were relevant for prudential regulation and include them in our 2007 dataset (a conservative assumption).

Table 3: **FSA Handbook (2007) provisions judged exclusively relevant for non-bank sectors**

Set of Handbook provisions	**Applies exclusively to...**
Credit Unions (CRED)	Credit unions
Prudential Sourcebook for Insurers (INSPRU)	Insurers
Prudential Sourcebook for Mortgage and Home Finance Firms, and Insurance Intermediaries (MIPRU)	Mortgage and home finance firms; insurance intermediaries
Prudential Sourcebook for UCITS Firms (UPRU)	UCITS firms
Interim Prudential Sourcebook for Friendly Societies (IPRU-FSOC)	Friendly societies
Interim Prudential Sourcebook for Insurers (IPRU-INS)	Insurers
Interim Prudential Sourcebook for Investment Businesses (IPRU-INV)	Investment businesses

Note: Sets of FSA Handbook (2007) provisions which apply exclusively to non-bank sectors are excluded from our 2007 dataset. Where sets of Handbook provisions combine provisions relevant for banks and provisions exclusively relevant for non-banks, we included the entire set of provisions in our 2007 dataset (e.g. Chapter 2 of the General Prudential Sourcebook (GENPRU 2)).

2.4 Cross-references for network analysis

Individual provisions often refer explicitly to other individual provisions or to sets of provisions. These "cross-references" can be within a single text (e.g. between individual CRR articles) or across texts (e.g. from a PRA rule to a CRR article).

We record explicit cross-references between individual provisions and their direction, and use them to represent our 2007 and 2017 datasets as networks. Regulatory text forms the nodes and cross-references between text forms the edges.

By only recording explicit cross-references between individual provisions, we exclude 'one-to-many' connections where an individual provision refers to an entire set of provisions. All else equal, this makes our measures a lower bound on interconnectivity within the prudential regulatory framework.

Table 4: **Comparison of different levels of data aggregation (2007 and 2017)**

Document	**Provision**	**Provision Example**	**Chapter**	**Chapter Example**	**Year**
Handbook	Rule	1.1.1	Chapter	BIPRU 1 Application	2007
Rulebook	Rule	1.1	Part	Fundamental Rules	2017
CRR	Article	Article 1	Title	Part One Title I	2017
EBA TS	Article	Article 1	Technical Standard Document	Regulation (EU) No 1423/2013	2017
Supervisory Statement	Paragraph	1.1	SS Document	SS3/13	2017

2.5 What is the relevant textual unit?

The text of the prudential regulatory framework has a hierarchical structure that aggregates individual provisions in increasingly broad topics. For example, CRR is structured hierarchically into Parts, Titles, Chapters, Sections and Articles; and key capital ratios are defined in *Part III (Capital Requirements), Title I (General Requirements, Valuation and Reporting), Chapter 1 (Required level of own funds), Section 1 (Own fund requirements for institutions), Article 92 (Own funds requirements)*.

We distinguish local versus global complexity with reference to individual provisions, which are natural units of observation that reflect the decision of the rule writer on what constitutes the most granular level of aggregation of legal language.

As an alternative to individual provision, we rely on aggregations used by the drafters of each document which delivers fairly homogeneous topics (see Table 4 for details). But different decisions for different documents will affect cross-document comparison (across both time and legal sources).

To explain how aggregation can affect our results, we focus on network measures. At provision-level, we use only cross-references, which have four main advantages. First, it is conservative in estimating the level of connectivity of the network, since it uses only provision-to-provision cross-references. Second, there is a high number of individual provisions, increasing the precision of statistical tests. Third, legal texts are clearly structured into individual provisions. Fourth, policy makers may prefer breaking down rules that become overly complex into several provisions, so that while individual provisions remain relatively simple, the overall structure becomes much complex.

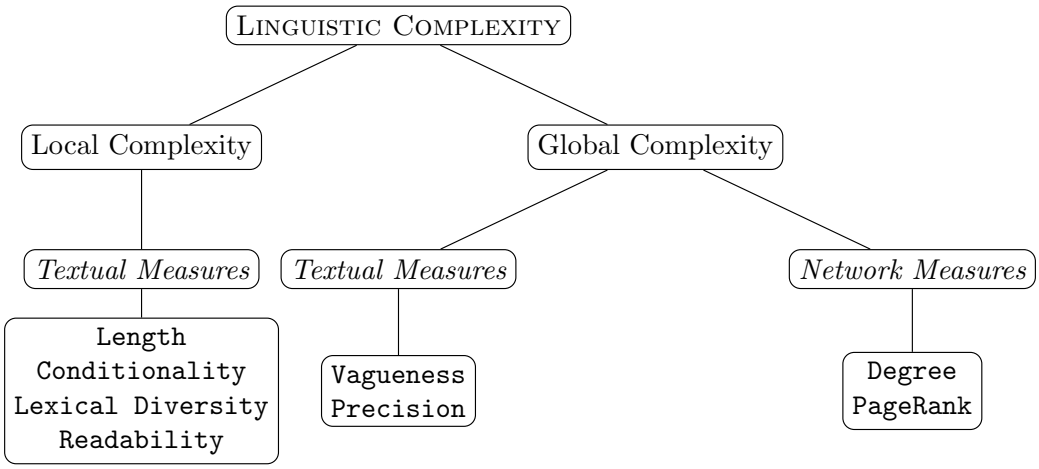
Using this level of aggregation has also disadvantages. First, we cannot use certain connections via hyperlinks between the PRA Rulebook and Supervisory Statements, and CRR and EBA Technical Standards, where hyperlinks often lead to whole documents. In such cases, it is dif-

ficult to attribute the connection to individual provisions. Second, to understand a provision it may not be sufficient to consider the provision itself, and follow the chain of cross-references. It may be necessary to take into account also other provisions on the same topic, even if not explicitly cross-referenced. To address these arguments, in section 4.4 we present results at a higher level of aggregation, which we define as "chapters".

3 Measures of linguistic complexity

In this section we provide first a motivation for using textual and network measures (Figure 3) to analyse the post-crisis banking reforms, then we describe our measures in detail.

Figure 3: **Measures of linguistic complexity**



3.1 Definition of complexity in the language sciences

Our definition of linguistic processing complexity draws on the notion of complexity in the cognitive psychology of language (psycholinguistics)—that is, the amount of processing difficulty encountered when producing or comprehending a particular linguistic unit (e.g. Gibson, 1998). Given that we will be largely taking the perspective of readers of provisions (language comprehenders) rather than that of the drafters (language producers), we will further limit our definition to language comprehension. A linguistic unit is considered to involve "processing difficulty" if it overtaxes the human comprehension processor within the brain, that is, if it impedes successful, fluent processing.[10]

To read a provision (sentence) successfully, what comprehension processing tasks does the reader have to accomplish? Upon encountering a word, the reader first has to accurately decode its orthography, retrieve its phonological code, access its lexical entry, syntactic category,

[10]Operationally, and using terms from eye-tracking methodology, a linguistic unit is complex if we *fixate* on it for longer relative to other linguistic units or *regress* back to the complex region.

and meaning from long-term memory. Having recognised the word, the reader next needs to integrate it into the sentence currently being processed, and update their present representation of the text, i.e. what has gone before. We do this incrementally as we proceed through the sentence, re-analysing if necessary, before performing sentence wrap-up.

Decades of research have shown that many things can impede (and facilitate) successful sentence processing (see Jaeger and Tily 2011 for a review). If a piece of text is too complex for whatever reason, it uses up too much of a reader's processing capacity, interferes with their working memory, and as a result makes it more difficult to integrate and encode new information into the current knowledge structure. Specifically, we know that two linguistic units that are structurally dependent are easier to integrate if they are adjacent compared with if they are separated (dependency length effect), because in the former case there are no other intervening linguistic units that need to be processed. Additionally, it is easier to retrieve subsequent mentions of a word that has currently been retrieved from memory, because it is already activated. Other factors include word length, frequency, and concreteness, with shorter, more frequent, and concrete words easier to process than longer, less frequent, and abstract ones. To process information successfully and overcome difficulties, the reader needs to not only recruit linguistic knowledge (especially syntax, semantics, and discourse inferencing) and information already in the text, but also external knowledge (plausibility, frequency). It may also be the case that a sentence is not fully comprehended until some additional information is processed (for instance, if we did not know and cannot infer the meaning of a word; needed to follow a link; or in crime-fiction novels had to wait 10 pages until we find out who the 'he' refers to in the first sentence). We store some representation of the sentence in memory, and then seek out the information that would be needed to fully resolve the partial understanding. If we don't already know what is being referenced, we have to go somewhere else to find it out and this additional step blocks successful on-line processing.

We take the liberty to adopt and adapt terms from the discourse processing literature, and categorize provision processing into two types depending on the site at which processing is likely to be complete. We use the term *local complexity* to refer to processing difficulties that are likely to be resolved while the reader is processing the text contained in the provision or during wrap-up. And we use the term *global complexity* to refer to processing difficulties encountered while reading that are likely to be resolved only after accessing information outside the immediate context of the provision—for instance, cross-references or regulatory precedents that help specify the meaning of some terms. We construct several measures, using natural language processing and network analysis to capture aspects of local- and global-level processing difficulties.

3.2 Local complexity measures

Our measures of local complexity are length, lexical diversity, conditionality, and readability:

Length At the provision-level, this measure attempts to capture the amount of linguistic material a reader has to retrieve, integrate, and temporarily store until sentence wrap-up is performed. There are many ways to computationally operationalise length, from number of words to perhaps more cognitively realistic measures such as the number of syntactic nodes and depth of syntactic embedding (see Szmrecsanyi (2004) for a review). Given that Szmrecsanyi (2004) finds all measures to be correlated, we use number of words (tokens) n as a quick way to approximate this concept.

Lexical Diversity Comprehension is facilitated when words are repeated. This effect is called repetition priming. A cognitively simple rule would thus have many repetitions (the same concept discussed over and over). A cognitively complex rule would have relatively little repetition (it would cover many different concepts). We operationalise this concept using a measure of *lexical diversity*, which measures the proportion of unique word types in a provision. This measure is usually computed as a type–token ratio, $Count(types)/Count(tokens)$, but as this measure is biased against longer documents, we use an adjustment using Maas' index (Maas, 1972):

$$LexicalDiversity_{Maas} = \frac{\log Count(tokens) - \log Count(types)}{\log Count(tokens)^2}. \tag{1}$$

Conditionality Sentences involving conditional clauses are complex in two ways: i) conditionals often deal with possible and counterfactual worlds, so we have to construct mental models of worlds that don't exist in order to be able to understand them (we can process facts better than hypothetical events), ii) if there are many different conditional clauses, we have to integrate many different exceptions, which may interfere with our ability to understand the applicability of a given provision. We measure conditionality by counting the number of conditional clauses (or conditional expressions) per sentence. We take the following words/phrases to indicate that conditionality is being introduced: *if, when(ever), where(ver), unless, notwithstanding, except, but, provided (that)*.

Readability The above features consider specific local-level complexity attributes and target specific aspects of processing. To get an overall impression of the complexity of a given rule at the local level, we use the Flesch-Kincaid grade level readability metric (Kincaid et al., 1975):

$$0.39 \times \frac{\#words}{\#sentences} + 11.8 \times \frac{\#syllables}{\#words} - 15.59. \tag{2}$$

3.3 Global complexity measures

Our measures of global processing complexity are based on work in network science (degree and PageRank) and NLP (vagueness and precision).

Degree and PageRank are based on provision-to-provision connections using cross-references. We summarise the regulatory framework as a network, where each provision is a node and each cross-reference is an edge between nodes. A provision (node) with more cross-references (edges) is more complex because a reader must leave the provision and visit other provisions/nodes to fully comprehend the provision itself.

Degree is one of the simplest descriptive statistics of a network. It describes a number of edges (or links) in a node, i.e. the number of connections from/to that node. Here, both incoming (in-degree) and outgoing (out-degree) edges are counted.

PageRank Page et al. (1999) summarises the centrality of a node within a network. Simplifying somewhat, PageRank counts the number and quality of cross-references *to* a provision to estimate how important the provision is. More important, i.e. more central, provisions are likely to get more direct and indirect links from other provisions. Degree only captures direct links, and, in this paper, ignores direction, while PageRank takes into account the whole web of indirect links that point towards a node. The two measures are likely to be correlated, e.g. in the case of a node that has many edges, both direct and indirect, that point towards it. But a node with many outgoing edges, and few incoming ones, will have high degree and low PageRank. And a node with only one incoming node will have low degree, but can have high PageRank, if the incoming node has, in turn, many incoming edges (direct and indirect). Degree is a more intuitive measure, but for the purposes of our analysis, PageRank helps capture "regulatory cascades"—chains of connected provisions. Changes to a provision whose corresponding node has high PageRank can have knock on effects on a chain of other provisions.

Our vagueness and precision measures use hand-crafted lexicons. Provisions that use only precise terms require little additional contextual information outside of the provision itself to be successfully processed. Numerical values are often used in legal texts to ensure that the text can be processed at "face value". Vague terms such as "reasonable" instead require additional information, not included in the specific provision, to be evaluated. Clarifying information could come from regulatory precedents, and in extreme cases may not be available at all.

Vagueness Our vagueness measure captures the extent to which the reader needs to use discretion and judgment in interpreting a given provision. Drawing on domain expertise, we create a lexicon on 22 vague terms and compute their frequency in each provision (normalising for provision length). A full list can be found in Table 5.

Table 5: **List of vague terms**

Term	
appropriat	possibl
adequate	potential
available	practicabl
effective	prompt
equitabl	reasonabl
fair	regular
good	several
likely	significant
material	substantial
necessary	sufficient
particular	timely

Precision Our precision metric assesses the rate of precise numerals in a given provision—specifically, amounts following indicators of currency ($£, \$, GBP, USD$, etc.) and per cents ($\%$).

3.4 What do local and global complexity measures tell us about regulatory complexity?

We now discuss how our measures of complexity, which are rooted in linguistics, relate to other concepts of complexity in the legal and economic literature on regulation.

The measures of local complexity try to proxy the number of eventualities (or states of the world, or instances) detailed in a provision (Ehrlich and Posner, 1974). Using a formal logic framework Battigalli and Maggi (2002) describe rules in terms of two components: primitive sentences that describe elementary events and tasks; and logical connectives such as *if* and *then*. Writing more complex rules requires both primitive sentences and more logical operators. Colliard and Georg (2018) draw instead on a parallel between rules and coding algorithms, and distinguish between logical "operators" and "operands", required to describe events and behaviours. We use lexical diversity to proxy unique primitive sentences or operands, and capture unique events and behaviours as unique concepts. Conditionality instead counts logical connectors or operands. Finally, using more primitive sentences (operands) and logical operators results in using more words, which we capture in our measure for length.

The measures of global complexity capture two different sources of contextual information: other provisions clearly cross-referenced in the prudential regulatory framework; and additional (unspecified) information that is likely to be necessary to interpret vague terms contained in the provision.

Cross-references reflect complexity that results from the structure, rather than the language, of rules. Ely (2011) and Kawai et al. (2018) stress that policy making is incremental, and can

lead to "kludging": policies are difficult to reform because they are entangled. Colliard and Georg (2018) point out that rule writers can simplify the algorithm by introducing "built-in" functions, that is intermediary results that define concepts that are then used in other rules. These "built-in" functions can be specialized terms (which would be capture by lexical diversity), but cross-references between provisions perform a similar role by reducing repetition. Our structural network measures capture entangling between provisions: the meaning of a rule with high PageRank (or degree) affects the meaning of other provisions that cross-reference to it (directly or indirectly via a chain of cross-references).

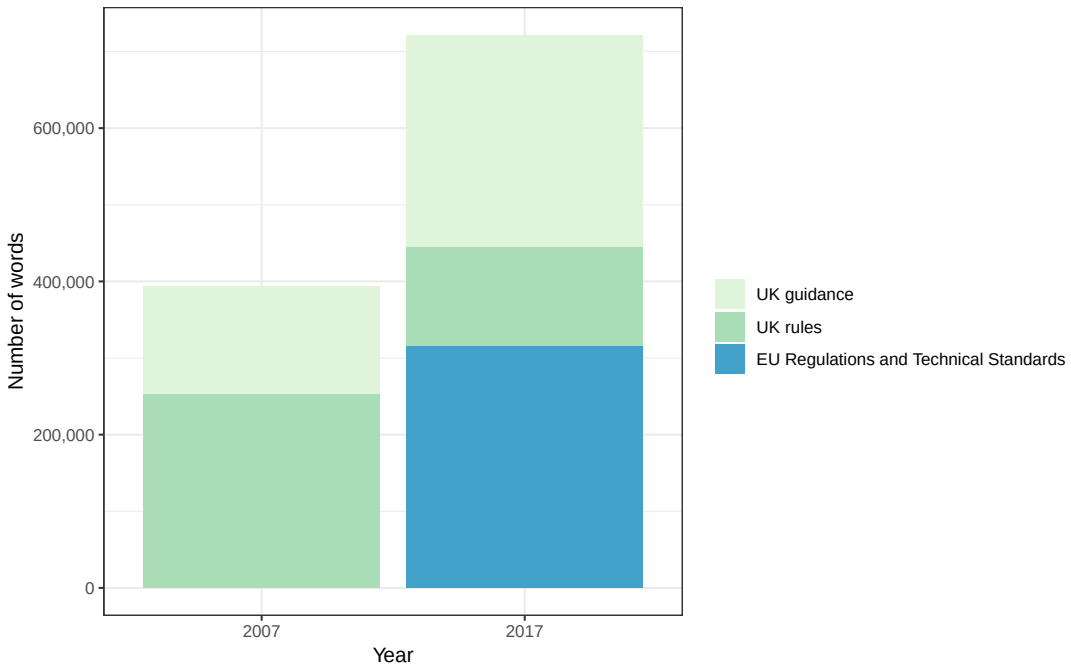
Finally, the "rules versus standards" literature focuses on the need to interpret the meaning of vague terms contained in a provision. In this literature a "rule" is a provision that contains precise terms, while a "standard" is a provision that achieves the same aim with vague terminology. A classic example (taken from Kaplow (1992)) is a speed limit expressed as "driving in excess of 55 miles per hour" (a rule) versus a limit expressed in terms of "driving at an excessive speed" (a standard).[11] Interpretation of vague terms requires accessing additional sources of information such as legal precedents and other sources of law (possibly with the intervention of a court). Our measures of vagueness and precision capture two extremes: on the one hand, extremely vague terms such as "reasonable" and "adequate'; on the other, numerical values.

3.5 Aggregate comparison and network maps

For the United Kingdom, we estimate that, in 2017, after most of the regulatory reforms were incorporated, the universe of prudential rules for banks included over 720,000 words, compared to almost 400,000 words in 2007, before the crisis (see Figure 4).

[11] In the context of this paper, the rules versus standard terminology is confusing (since we define rules as legally binding provisions) and we avoid it.

Figure 4: **Length of regulatory prudential framework for UK deposit-takers (2007 and 2017)**



Note:
UK guidance: PRA Supervisory Statements (2017), guidance in FSA Handbook (prudential guidance only, 2007).
UK rules (including rules implementing EU Directives): PRA Rulebook (2017), FSA Handbook (prudential rules only, 2007).
EU Regulations and Technical Standards: Capital Requirements Regulation (2013) and related EBA Technical Standards.

Table 6: **Comparison of descriptive statistics (2007 and 2017)**

	2007	2017	Change %
Panel A: Number of words			
Total	393,290	721,642	83.5%
Rules	221,912	445,710	100.8%
Guidance	171,378	275,932	61.0%
Panel B: Unique words			
Total	11,799	13,420	13.7%
Rules	8,657	9,671	11.7%
Guidance	7,558	8,765	16.0%
Panel C: Conditional words			
Total	3,930	6,320	60.8%
Rules	2,250	3,802	69.0%
Guidance	1,680	2,518	49.9%
Panel D: Provision-level network (rules only)			
Nodes (Provisions)	2,381	2,482	4.2%
Edges (Cross-references)	2,569	4,289	67.0%
Panel E: Provision-level network (rules and guidance)			
Nodes (Provisions)	4,439	5,142	15.8%
Edges (Cross-references)	5,717	4,809	-15.9%
Panel F: Chapter-level network (rules and guidance)			
Nodes (Chapters)	80	145	81.2%
Edges (Cross-references)	1038	2131	105.3%

Table 6 compares aggregate figures for the prudential regulatory framework for UK banks in 2007 and 2017.[12] Panel A shows that the total number of words increased from almost 400,000 to over 720,000 (+83%). Rules account for the majority of words (445,000), with the number of words in rules doubling between 2007 and 2017. Words in guidance also increased substantially (+61%).

Panel B shows that the number of unique words increased from almost 12,000 in 2007 to over 13,000 in 2017 (+14%). The number of unique words did not increase proportionally to the number of words. This is not surprising, since the likelihood of repeating words increases with the volume of words. Interestingly, guidance accounts for almost half of the unique words in 2017, even if it represents a lower share of total words.

Panel C shows that the number of conditional expressions increased from almost 4,000 to over 6,000 (+61%), with similar changes for rules and guidance.

Panels D through F compare the numbers of network nodes and edges in 2007 and 2017. In Panel D, nodes correspond to individual provisions and edges to cross-references between provisions. The number of nodes/provisions increases slightly from almost 2,400 to almost 2,500 (+4%), but the change in the number of edges/cross-references is more substantial, from about 2,500 to almost 4,300 (+67%). This result indicates that, when we consider only rules, the provision-level network becomes more interconnected between 2007 and 2017.

[12] We do not include headers and footers.

Panel E includes both rules and guidance, at provision level. The number of edges/cross-references decreases over 5,700 to 4,800 (-16%), while the number of nodes/provisions grows from over 4,400 to over 5,100 (+15%). The fall in edges is driven by guidance, and could be explained by the transition from the FSA to the PRA. While FSA guidance was included in the Handbook and drafted by lawyers, PRA guidance is published in separate documents (Supervisory Statements) and is no longer drafted by lawyers. Many provisions have no edges in 2017. The proportion of these isolated nodes increased from 31% to 42% of the total number between 2007 and 2017.

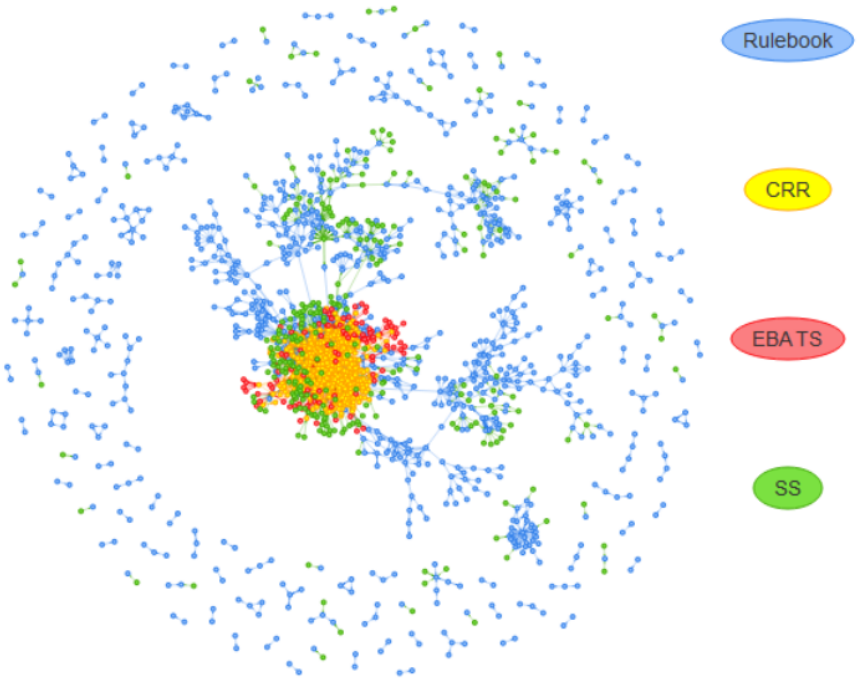
Panel F presents figures for the network at chapter level (rules and guidance combined). Each node is a chapter, and each edge a cross-reference between provisions in different chapters. Given that chapters aggregate several provisions, both the number of nodes and the number of edges are lower.[13] Between 2007 and 2017, the number of nodes increases from 80 to 145 (+81%), and the number of edges increases from 1,038 to 2,131 (+105%).

Figure 5 provides a visualisation of the network for 2007. Each point (node) represents a provision, and each edge (line between nodes) is a cross-reference. Figure 6 shows the network for 2017.[14] For 2017, we also show the legal source for each provision: CRR (yellow nodes), Rulebook (blue), EBA TS (red) and Supervisory Statements (green). For visual clarity, only nodes with at least one edge are displayed. A visual comparison between the two figures highlights how the 2017 has a denser core, but also a larger periphery of nodes with only one of two edges. CRR provisions are at the centre of the network.

[13] Edges between provisions in the same chapter are not included in calculations at chapter level.
[14] Edges on the plot are unweighted, i.e. they do not show the number of cross-references.

Figure 5: **Provision-level network (2007)**

Figure 6: **Provision-level network (2017)**



4 Results

In this section we present our results. We discuss first results for local (Section 4.1) and then global complexity (Section 4.2), both using individual provisions as the relevant textual unit. We also compare the complexity of legal sources in 2017 (in Section 4.3). We conclude the section by showing how our results change when we use a higher level of aggregation (Section 4.4).

4.1 Local complexity

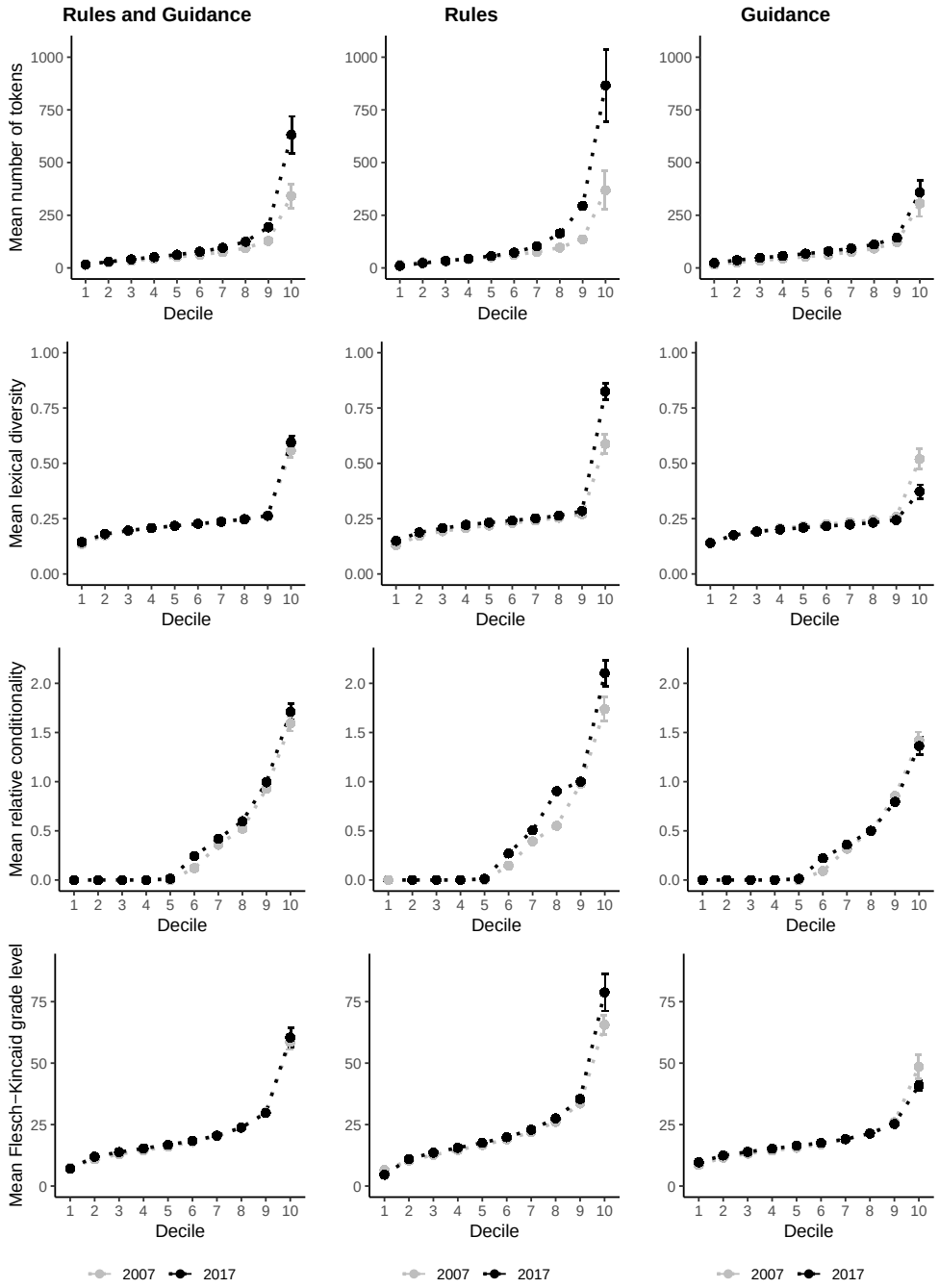
Our local complexity measures capture comprehension of individual provisions *while* they are being read, abstracting from other sources of information that the reader may need to access at a later stage. Figure 7 compares the distributions of our first set of measures in 2007 (grey) versus 2017 (black) on provision-level. Each row corresponds to one of the measures. The charts in the first column show results for all provisions (both rules and guidance combined), while the second and third columns show results for rules only and guidance only, respectively.

To construct the decile plots, we calculate the relevant measure of complexity for each provision. Provisions in each year are then ordered by the relevant measure and then split into ten bins (deciles). We display the mean for each bin. For example, we calculate the length of each provision in our dataset for 2017. We then rank each of the provisions in 2017 terms of length, create ten decile bins, and calculate the average length within each bin. We repeat the process for 2007. Finally, we plot the average for each decile in 2007 and 2017, and compare the plots.

Length When we consider all prudential provisions (first row, first column), the mean of tokens (words) is similar in 2007 and 2017 for the bins in the bottom half of the distribution. Provisions for 2017 become longer in the upper half of the distribution, and a very large gap is evident for provisions in the top decile. This pattern is more pronounced when we look at rules only. The provisions in the the top end of the distribution for guidance are not particularly long, and they become only marginally longer in 2017.

Lexical diversity The distributions for lexical diversity appear virtually identical in 2007 and 2017 when we include all provisions (second row), but this masks opposite trends for rules and guidance. Rules provisions in the top bin are more lexically diverse in 2017 , while guidance provisions in the top bin are less diverse.

Figure 7: **Linguistic measures on provision-level by regulation type (2007 and 2017)**



Relative conditionality For relative conditionality (third row), the distributions for all provisions are similar for 2007 and 2017. For rules only, mean conditionality is higher in 2017 in the top half of the distributions. For guidance only, conditionality is slightly lower in the distributions.

Readability Finally, the distributions for the Flesch-Kincaid grade level scores show a decline in 2017 in the highest deciles, driven by guidance, while rules have slightly higher scores in 2017.

To summarise, two observations emerge from the analysis of our textual measures of local information processing. First, the change between 2007 and 2017 is concentrated in the top deciles, while there is very little movement at the low end, but also the middle, of the distributions. We do not observe a shift of the whole distribution for any of the measures. This result seems to indicates a continuity in the style of writing provisions over time, suggesting that the authors try to limit the complexity of individual provisions. Authors may decide to break a provision that is particularly long or contains several different concepts (captured by lexical diversity) into two or more different provisions.

Second, we see diverging patterns in rules versus guidance. Compared to 2007, the top deciles of rule provisions in 2017 contain longer provisions, with more lexical diversity, and conditionality. This indicates that post-crisis reforms have generated a (small) set of rules that require retrieving a lot of material, containing several concepts and exceptions. Guidance, on the other hand, has changed less as a result of the reforms. The main change is a reduction in lexical diversity in the top bin, indicating that some guidance provisions that contained a large range of concepts may have disappeared.

4.2 Global complexity

In this section, we present results for our measures of global processing required for comprehension. The first two measures, degree and PageRank, capture the network structure of the regulatory framework using cross-references between provisions, while the last two, vagueness and precision, are linguistic measures.

4.2.1 Network measures

Measures of network structure in Figure 8 summarise changes in interconnection between provisions via cross-references.

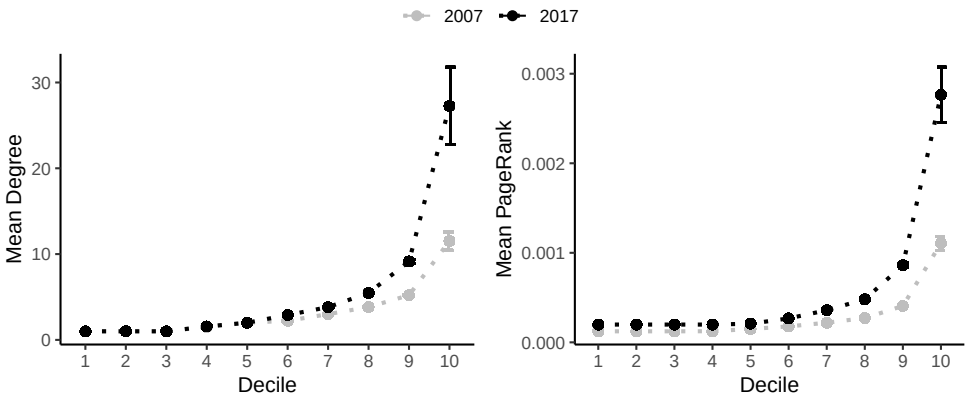
Degree The left-hand side chart shows the distributions for degree. In the lower half of the distribution, the difference in degree between 2007 and 2017 appears small. A wedge appears

in the upper half of the distribution. The nodes in the top bin (top ten per cent) in 2017 have a mean degree of about 30 in 2017, almost three times as high as in 2007.

PageRank Similarly, the chart on the right-hand side of Figure 8 plots the 2007 and 2017 distributions for PageRank. The difference is concentrated in the top bin, where mean PageRank is also about three times higher in 2017 compared to 2007 (but the absolute value of PageRank does not have an intuitive interpretation).

Figure 8 also shows 95% confidence intervals at each bin, but for several bins the coefficients for 2007 and 2017 appear to be both very close and precisely estimated.

Figure 8: **Network centrality measures on provision-level by year (2007 and 2017)**



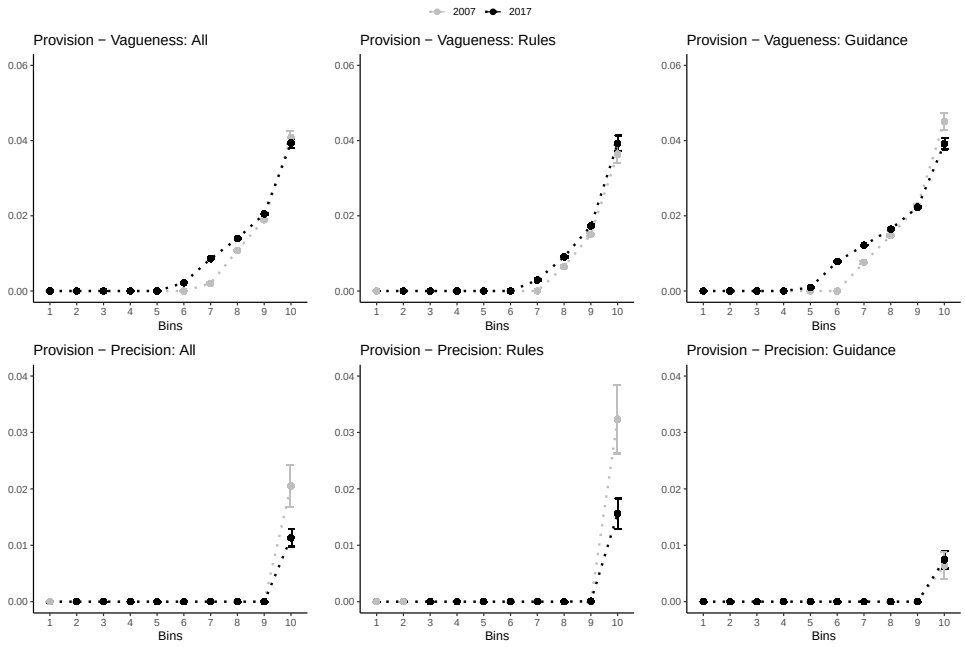
4.2.2 Vagueness and precision

Figure 9 compares distributions in 2007 and 2017 and presents results for our vagueness and precision measures. As for other linguistic measures we show distributions for all provisions (both rules and guidance), rules only, and guidance only.

Vagueness We start with our vagueness measure, which simply picks the proportion of vague terms in the text of a provision. About half of all provisions do not contain any of our vague terms, but in the upper part of the distribution we see more vague terms in 2017. For rules only, we see a parallel shift upwards in the upper part of the distribution. For guidance only, the increase is concentrated in the middle bins, while there is a decrease in vague terms in top bins.

Precision Only provisions in the top decile contain any of the precise terms (currency and per cent), and the proportion almost halves in 2017 compared to 2007. This is driven by rules, which contain a higher proportion of precise terms than guidance.

Figure 9: **Vagueness by regulation type and year (2007 and 2017)**



4.3 Sources

In this section, we compare our textual measures across source documents. We concentrate on provisions for 2017 from four sources: CRR, EBA Technical Standards, PRA Rulebook, and PRA Supervisory Statements.

Figure 10 compares the distributions of different source documents for our four measures of local complexity. Provisions in CRR and EBA TS have higher numbers of tokens (i.e. are longer) in the upper part of the distribution, while the distributions for PRA Rulebook and PRA Supervisory Statements overlap. EBA TS also have high Flesch-Kincaid scores (low readability) in the top bin. Provisions in the top bin for the distribution for the PRA Rulebook have very high lexical diversity and high relative conditionality (compared to other sources). The distribution for PRA Supervisory Statements is below those for other sources across all measures.

Figure 11 plots degree and PageRank distributions by document source. Both degree and PageRank distributions are higher for CRR than for EBA TS or PRA Rulebook. The degree distributions for EBA TS and Rulebook are very similar, except in the top decile, where mean degree for EBA TS increases (but is imprecisely estimated). The opposite happens for PageRank.[15]

[15]EBA Technical Standards implement the details of CRR articles, and understanding how to comply with a specific CRR article requires reading also the related EBA TS. The CRR articles however do not contain cross-references to specific articles in EBA Technical Standards (which had not been written yet), but EBA TS contain cross-references to CRR articles. The online version of CRR contains hyper-links to the relevant implementing document for EBA TS, not individual articles. To remain faithful to the textual structure of the document, we only retain cross-references from EBA TS to CRR. EBA TS articles have outgoing edges, but almost no incoming edges, and as a result can have high degree, but not high PageRank.

Figure 10: **Linguistic measures on provision-level by source (2017)**

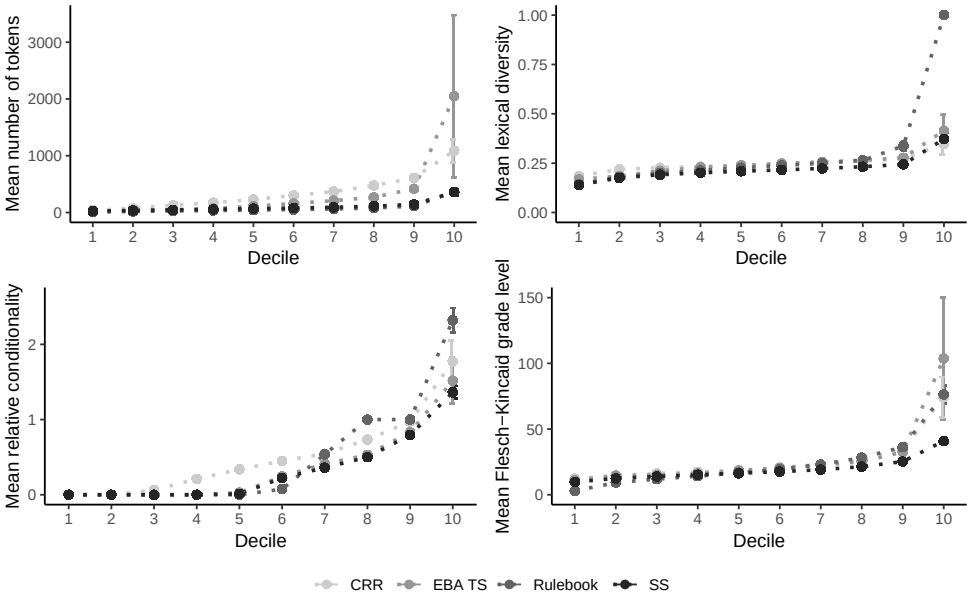
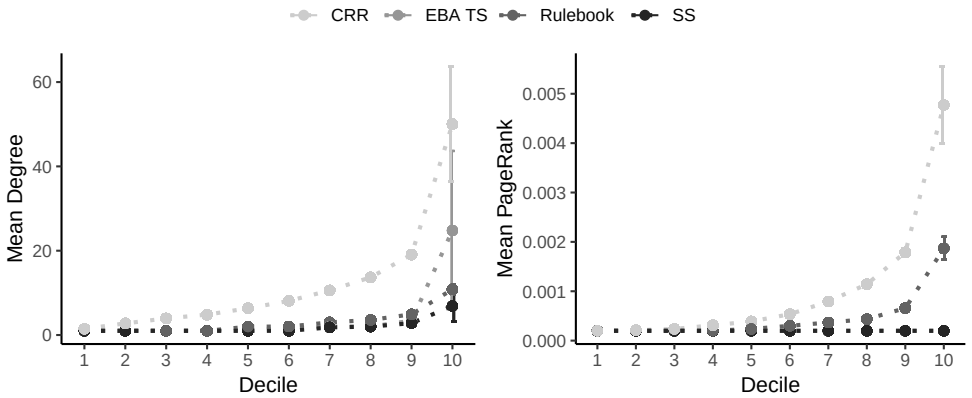


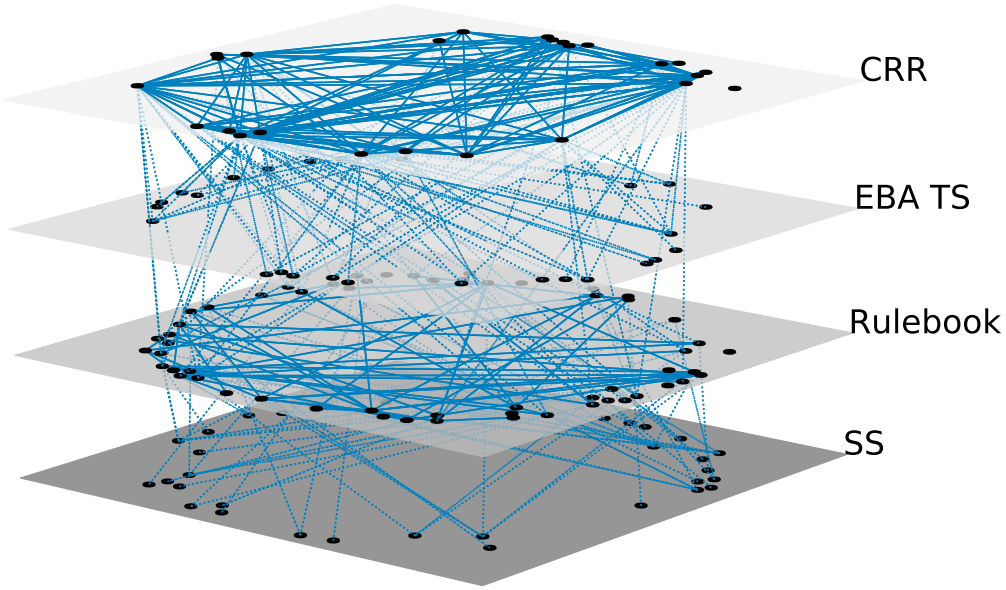
Figure 11: **Network centrality measures on provision-level by source (2017)**



4.4 Chapter-level results

In this section we present results using a higher level of aggregation as textual unit—a "chapter" as defined in Table 4. A graphical representation of connections between different documents is presented in Figure 12.

Figure 12: **Chapter-level multilayer network (2017)**



Note: Links from Supervisory Statements to the Rulebook.

Local complexity Figure 13 shows distributions for linguistic measures at chapter level. For ease of comparison, Figure 13 is structured in the same way as Figure 7, which displays linguistic measures at provision-level. Each row corresponds to a measure, while the three columns show results for all chapters (rules and guidance), rules only and guidance only. Given the low number of observations we do not display confidence intervals.

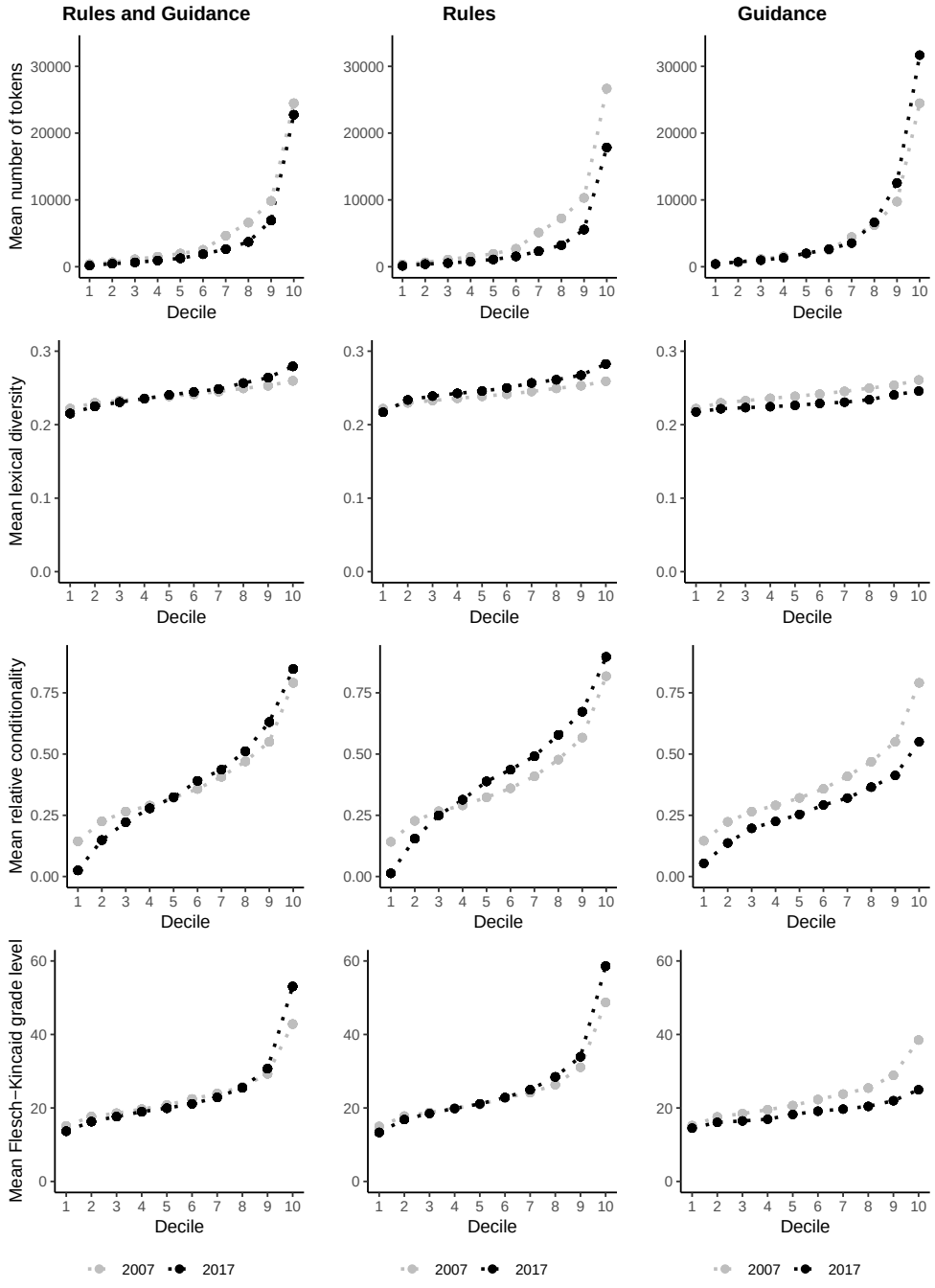
The mean number of words is obviously much higher at chapter level, but it is lower for 2017 than for 2007, in particular for rules. This reflects the composition of the source documents used in 2017 vs. 2007, and their different structure.

For lexical diversity, the results are qualitatively similar to those at provision-level. The increase in lexical diversity is driven by rules and concentrated in the top half of the distribution. Lexical diversity for guidance instead is lower in 2007. However, the change between 2007 and 2017 is much more visible. This can be in part due to lower precision of the estimates due to the much smaller sample at chapter level. But it can also indicate that the addition of new concepts as a result of the post crisis reforms has led to higher lexical diversity within chapters. The change is not visible at provision level suggesting that additional concepts are distributed over different provisions.

Results for relative conditionality at chapter level are also similar to those at provision level, with opposite direction of change for rules versus guidance, and these changes are more visible and not restricted to the top deciles.

For the Flesch-Kincaid readability scores, the reduction for guidance is much more visible at chapter- than provision-level.

Figure 13: **Linguistic measures on chapter-level by regulation type (2007 and 2017)**



Global complexity Figure 14 shows degree and PageRank distributions in 2007 and 2017. Compared to the analogous provision-level distributions in Figure 8, the order of magnitude of interconnections is much higher, e.g. the top bin of the degree distribution in 2017 contains on average 300 links at chapter level, compared to about 30 links at provision level. The PageRank distribution for 2017 is above the 2007 distribution only in the top decile, and Figure 15 shows that almost all the variation can be attributed to CRR.

Figure 14: **Network centrality measures on chapter-level by year (2007 and 2017)**

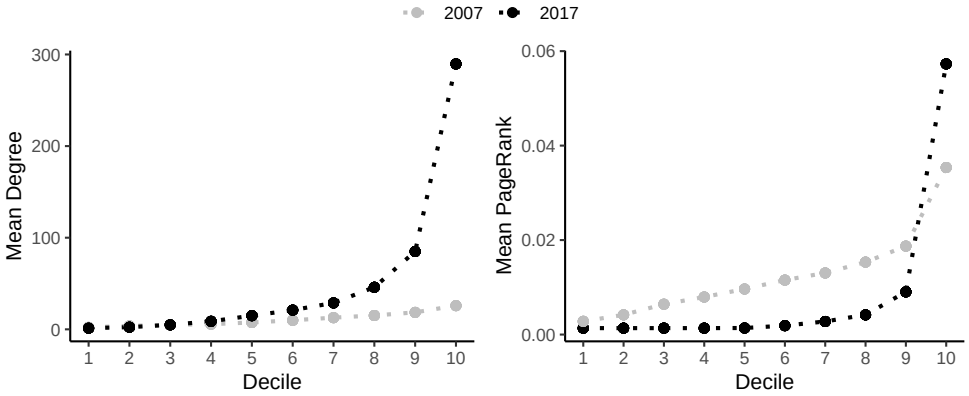
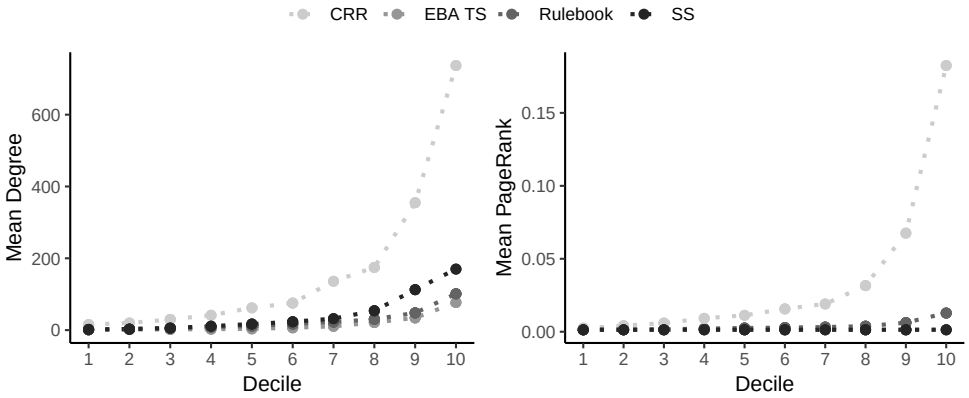


Figure 15: **Network centrality measures on chapter-level by source (2017)**



5 Conclusions

Post-crisis bank regulation reforms expanded to fix the flaws exposed by the crisis. In the EU, Single Rulebook harmonisation resulted in a patchwork of source documents across European Commission, EU supervisory authorities, and national regulators (Bassani, 2019). In the same period, the growth of artificial intelligence techniques has opened the prospect of machine-aided or "cyborg" supervision of banks (Proudman, 2018).

These two developments have created new sets of issues for bank regulators. One relates to the effectiveness of post-crisis bank policy. Regulators are now moving into an evaluation phase and starting to question whether interactions between individual regulatory policies—designed and implemented separately in response to the crisis—are creating unintended consequences that detract from policy intent. But without empirical measures of sources and directions of interconnections within the regulatory system, it is hard to say much about exactly and where the unintended consequences might be, or to go beyond normative judgments when thinking about the system as a whole. The results in this paper suggest that the interconnectivity in the framework stems from a relatively small set of provisions, which can be identified using our measures.

A separate but related set issue derives from the growing interest in the potential for machine-readable rules to increase regulatory efficiency and reduce compliance costs (Casey and Niblett, 2016). But the net benefits from automating the reading and execution of rules will vary across rule types: machines may assimilate long, highly conditional rules more effectively than humans can, but are less well suited for assimilating vaguer, more context-specific provisions. Further research could focus on developing the measure of vagueness presented in this paper to help identify the scope for automating the execution of rules.

The regulatory framework for banks includes both rule types to facilitate efficiency and predictability, and flexibility and supervisory discretion, when required. A regulatory framework that is too predictable ("tick-box"), with mechanical application of the rules, may be more efficient in an artificial intelligent environment, but it may be also be insufficiently adaptable to deal with creative compliance and even contribute to the build-up of risks (Danielsson et al., 2017). Ultimately, the question is how to develop a model of financial regulation that is both scalable, so that it can deal with the complexity of financial markets, and responsive to changes in the behaviour of regulated firms (Ford, 2013).

The ultimate normative question is "How complex does bank regulation have to be?" Regulators seek to control a highly complex system, with multiple interconnected agents, and large externalities. Because this system is adaptive—e.g. we observe herd behaviour and regulatory arbitrage—a regulatory system of individually complex and interacting parts could become less effective over time than a simpler one at mitigating financial stability risks (Aikman et al., 2018). But before we answer the normative question, we need to answer a positive one: "How complex is bank regulation?" In this paper, we have provided evidence which helps us answer

this positive question by calculating textual complexity indicators on the near universe of UK prudential rules. We have released key elements of our code alongside this paper and will release more at a later stage.

We recognise that more benchmarking for these indicators is a necessary next step towards answering the normative question. Benchmarking against non-financial regulatory frameworks, or frameworks in other jurisdictions, is challenging given differences in legal systems and policy substance. But we plan to exploit variation within our dataset to compare changes in complexity measures for different policy standards and test how they correspond to policymakers' expectations. We stress that these measures do not exhaust all the dimensions of linguistic complexity—in particular, resolving ambiguity in regulation is very likely to be important for information burden. In addition, to understand the economic effect of regulatory complexity "soft" textual information needs to be combined with traditional "hard" numeric data. For example, textual regulatory complexity could be compared to balance sheet complexity captured in Cetorelli and Goldberg (2014) and Goldberg and Meehl (2019). Eventually, NLP can help enrich the economic evaluation of rules in terms of: 1) the interaction between rules; 2) the impact of linguistic complexity; 3) the effectiveness of "rules vs standards." We contribute to this long-term research agenda, by creating a dataset of all provisions for UK banks, and analysing how they have changed with post-crisis reforms.

References

Aikman, D., M. Galesic, G. Gigerenzer, S. Kapadia, K. V. Katsikopoulos, A. Kothiyal, E. Murphy, and T. Neumann (2014). Taking uncertainty seriously: simplicity versus complexity in financial regulation. *Bank of England Financial Stability Paper* (28).

Aikman, D., A. Haldane, M. Hinterschweiger, and S. Kapadia (2018). Rethinking financial stability. *Bank of England Staff Working Paper* (712).

Ash, E. (2015). The political economy of tax laws in the US states. *Working Paper, Columbia University*.

Bardoscia, M., S. Battiston, F. Caccioli, and G. Caldarelli (2017). Pathways towards instability in financial networks. *Nature Communications 8*, 14416.

Bassani, G. (2019). *The Legal Framework Applicable to the Single Supervisory Mechanism: Tapestry or Patchwork?* Kluwer Law International BV.

Battigalli, P. and G. Maggi (2002). Rigidity, discretion, and the costs of writing contracts. *American Economic Review 92*(4), 798–817.

Battiston, S., J. D. Farmer, A. Flache, D. Garlaschelli, A. G. Haldane, H. Heesterbeek, C. Hommes, C. Jaeger, R. May, and M. Scheffer (2016). Complexity theory and financial regulation. *Science 351*(6275), 818–819.

Battiston, S., M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli (2012). DebtRank: Too central to fail? financial networks, the FED and systemic risk. *Scientific reports 2*, 541.

Behn, M., R. Haselmann, and P. Wachtel (2016). Procyclical capital regulation and lending. *The Journal of Finance 71*(2), 919–956.

Bholat, D., S. Hansen, P. Santos, and C. Schonhardt-Bailey (2015). Text mining for central banks. *Available at SSRN 2624811*.

Black, J. (1997). *Rules and regulators*. Clarendon Press Oxford.

Black, J. (2014). Learning from regulatory disasters. *Policy Quarterly 10*.

Born, B., M. Ehrmann, and M. Fratzscher (2014). Central bank communication on financial stability. *The Economic Journal 124*(577), 701–734.

Brochet, F., P. Naranjo, and G. Yu (2012). Causes and consequences of linguistic complexity in non-us firm conference calls. *Harvard Business School working paper series #13-033*.

Buehlmaier, M. M. and T. M. Whited (2018). Are financial constraints priced? evidence from textual analysis. *The Review of Financial Studies 31*(7), 2693–2728.

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.

Casey, A. J. and A. Niblett (2016). The death of rules and standards. *Ind. LJ 92*, 1401.

Cetorelli, N. and L. S. Goldberg (2014). Measures of global bank complexity. *FRBNY Economic Policy Review 20*(2), 107–126.

Chakraborty, C. and A. Joseph (2017). Machine learning at central banks. *Bank of England Staff Working Paper* (674).

Chen, H., P. De, Y. J. Hu, and B.-H. Hwang (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies 27*(5), 1367–1403.

Claessens, S. et al. (2017). *Regulation and structural change in financial systems*. Centre for Economic Policy Research.

Colliard, J.-E. and C.-P. Georg (2018). Measuring regulatory complexity. *Working Paper*.

Cooter, R. and T. Ulen (2016). *Law and economics*. Addison-Wesley.

Da, Z., J. Engelberg, and P. Gao (2014). The sum of all fears investor sentiment and asset prices. *The Review of Financial Studies 28*(1), 1–32.

Danielsson, J., R. Macrae, and A. Uthemann (2017). Artificial intelligence, financial risk management and systemic risk.

de Souza, J. A. S., J. C. Rissatti, S. Rover, and J. A. Borba (2019). The linguistic complexities of narrative accounting disclosure on financial statements: An analysis based on readability characteristics. *Research in International Business and Finance 48*, 59–74.

Duffie, D. (2017). Financial regulatory reform after the crisis: An assessment. *Management Science 64*(10), 4835–4857.

Ehrlich, I. and R. A. Posner (1974). An economic analysis of legal rulemaking. *The Journal of Legal Studies 3*(1), 257–286.

Ely, J. C. (2011). Kludged. *American Economic Journal: Microeconomics 3*(3), 210–31.

Endicott, T. (2005). The value of vagueness. *Vagueness in normative texts 23*, 27–48.

Engelberg, J. E., A. V. Reed, and M. C. Ringgenberg (2012). How are shorts informed?: Short sellers, news, and information processing. *Journal of Financial Economics 105*(2), 260–278.

Enriques, L., A. Romano, and T. Wetzer (2019). Network-sensitive financial regulation. *The Journal of Corporation Law*.

Ford, C. (2013). Prospects for scalability: Relationships and uncertainty in responsive regulation. *Regulation & Governance 7*(1), 14–29.

Fowler, J. H., T. R. Johnson, J. F. Spriggs, S. Jeon, and P. J. Wahlbeck (2007). Network analysis and the law: Measuring the legal importance of precedents at the US Supreme Court. *Political Analysis 15*(3), 324–346.

Gai, P., M. Kemp, A. Sánchez Serrano, and I. Schnabel (2019). Regulatory complexity and the quest for robust regulation. Technical report, European Systemic Risk Board.

Galati, G. and R. Moessner (2018). What do we know about the effects of macroprudential policy? *Economica 85*(340), 735–770.

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance 68*(3), 1267–1300.

Gentzkow, M., B. T. Kelly, and M. Taddy (2017). Text as data. Technical report, National Bureau of Economic Research.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition 68*(1), 1–76.

Gigerenzer, G. and R. Selten (2002). *Bounded rationality: The adaptive toolbox*. MIT press.

Glaeser, E. L. and A. Shleifer (2002). Legal origins. *The Quarterly Journal of Economics 117*(4), 1193–1229.

Goldberg, L. S. and A. Meehl (2019). Complexity in large US banks. *FRB of New York Staff Report* (880).

Greenwood, R., J. C. Stein, S. G. Hanson, and A. Sunderam (2017). Strengthening and streamlining bank capital regulation. *Brookings Papers on Economic Activity 2017*(2), 479–565.

Gropp, R., T. Mosk, S. Ongena, and C. Wix (2018). Banks' response to higher capital requirements: Evidence from a quasi-natural experiment. *The Review of Financial Studies 32*(1), 266–299.

Guay, W., D. Samuels, and D. Taylor (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics 62*(2-3), 234–269.

Hakenes, H. and I. Schnabel (2014). Regulatory capture by sophistication. *CEPR Discussion Paper No. DP10100*.

Haldane, A. (2017). A little more conversation, a little less action. *Bank of England - Speech*.

Haldane, A. G. and V. Madouros (2012). The dog and the frisbee. *Speech at the Federal Reserve Bank of Kansas City's 36th economic policy symposium, Jackson Hole, Wyoming*.

Handley, K. and J. F. Li (2018). Measuring the effects of firm uncertainty on economic activity: New evidence from one million documents. Technical report, Mimeo., University of Michigan.

Hoberg, G., G. Phillips, and N. Prabhala (2014). Product market threats, payouts, and financial flexibility. *The Journal of Finance 69*(1), 293–324.

Jaeger, T. F. and H. Tily (2011). On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science 2*(3), 323–335.

Juelsrud, R. E. and E. G. Wold (2018). Risk-weighted capital requirements and portfolio rebalancing. *Journal of Financial Intermediation*.

Kaplow, L. (1992). Rules versus standards: An economic analysis. *Duke Lj 42*, 557.

Kaplow, L. (1995). A model of the optimal complexity of legal rules. *JL Econ. & Org. 11*, 150.

Katz, D. M. and M. J. Bommarito (2014). Measuring the complexity of the law: the United States Code. *Artificial intelligence and law 22*(4), 337–374.

Kawai, K., R. Lang, and H. Li (2018). Political kludges. *American Economic Journal: Microeconomics 10*(4), 131–58.

Kincaid, J. P., R. P. Fishburne Jr, R. L. Rogers, and B. S. Chissom (1975). Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch reading ease formula) for navy enlisted personnel.

Koniaris, M., I. Anagnostopoulos, and Y. Vassiliou (2017). Network analysis in the legal domain: A complex model for European Union legal sources. *Journal of Complex Networks 6*(2), 243–268.

Kroll, J. A., S. Barocas, E. W. Felten, J. R. Reidenberg, D. G. Robinson, and H. Yu (2016). Accountable algorithms. *U. Pa. L. Rev. 165*, 633.

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.

Levmore, S. and F. Fagan (2019). The impact of artificial intelligence on rules, standards, and judicial discretion. *University of Chicago, Public Law Working Paper* (704).

Li, W., P. Azar, D. Larochelle, P. Hill, and A. W. Lo (2015). Law is code: a software engineering approach to analyzing the United States Code. *Journal of Business & Technology Law 10*, 297.

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

Loughran, T. and B. McDonald (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research 54*(4), 1187–1230.

Maas, H.-D. (1972). Über den zusammenhang zwischen wortschatzumfang und länge eines textes. *Zeitschrift für Literaturwissenschaft und Linguistik 2*(8), 73.

Page, L., S. Brin, R. Motwani, and T. Winograd (1999). The PageRank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab.

Perrow, C. (2011). *Normal accidents: Living with high risk technologies (Updated edition)*. Princeton university press.

Posner, R. A. (1975). The social costs of monopoly and regulation. *Journal of Political Economy 83*(4), 807–827.

Proudman, J. (2018). Cyborg supervision – the application of advanced analytics in prudential supervision. *Bank of England - Speech*.

Romano, R. (2012). Regulating in the dark. Volume 86, pp. 95–106. University of Pennsylvania Press Pennsylvania.

Ruhl, J. B. and J. Salzman (2002). Mozart and the red queen: The problem of regulatory accretion in the administrative state. *Geo. LJ 91*, 757.

Simon, H. A. (1991). The architecture of complexity. In *Facets of systems science*, pp. 457–476. Springer.

Szmrecsanyi, B. (2004). On operationalizing syntactic complexity. In *Le poids des mots. Proceedings of the 7th international conference on textual data statistical analysis. Louvain-la-Neuve*, Volume 2, pp. 1032–1039.

Tarullo, D. K. (2019). Financial regulation: Still unsettled a decade after the crisis. *The Journal of Economic Perspectives 33*(1), 61–80.