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

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 implementation of Basel III 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. These reforms have also led to concerns about the complexity of financial regulation. However, the debate lacks clarity on: (1) how to measure this complexity; and (2) the extent to which technology can be used to address it. We restrict our analysis to cognitive costs related to language processing, and construct a new textual dataset of the prudential rules that applied to UK banks before and after the Basel reforms. We use natural language processing and network analysis to calculate complexity measures on this novel dataset. We find that, while the language of individual rules remained stable, suggesting a continuity in drafting style, rules became more interconnected, via longer chains of cross-references. In following these chains, the number of words that a reader had to process starting from a single rule increased from about 600 words to over 25,000 on average (an increase of over 4,000%). We also contribute to developing textual measures that can help identify rules that are complex for humans but require limited interpretation and are better suited for translation into machine-readable code.

---

(1) PS AI Labs. Work conducted while at the Bank of England. Email: Zahid.Amadxarif@psail.com
(2) Bank of England. Email: James.Brookes@bankofengland.co.uk
(3) Ifo Institute. Work conducted while at the Bank of England. Email: Garbarino@ifo.de
(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 Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

1 Introduction

The Basel III reforms, a profound revision of international bank rules that followed the 2008 financial crisis, led to concerns about the complexity of the new regulatory framework (Romano, 2012; Haldane and Madouros, 2012; Greenwood et al., 2017). High cognitive costs of regulation can confuse both those who have to comply with rules (Feldman et al., 2016), and those who have to implement them (Ely, 2011). The cognitive efforts required to process (read and understand) regulatory texts are not the only source of regulatory costs, but they affect how firms comply and respond to regulation. Complex "red tape" can have implications for competition between incumbents and new entrants (Hakenes and Schnabel, 2014; Gutiérrez and Philippon, 2019). Traditional recommendations to simplify rules (Epstein, 1997; Sunstein, 2013) may however be less relevant if machines can process regulatory inputs (Casey and Niblett, 2016). To accelerate the adoption of machine-readable rules, international regulators have launched a series of "tech sprints" (Bank for International Settlements, 2020).

Cognitive costs are difficult to measure outside of experimental environments, and, for regulation, they are often roughly approximated using the length of the rulebook. As a result, the debate on the complexity of financial regulation, including Basel III reforms, often lacks clarity about what exactly is meant by "complexity" (Colliard and Georg, 2020). We restrict our analysis to cognitive costs related to language processing, and calculate established measures of textual complexity derived from network science, linguistics and legal studies.

To understand how Basel III affected these measures, we compare regulations that applied to banks operating in the United Kingdom in 2007 and 2017. The Basel III package was agreed internationally in 2010, and implemented in EU and UK law through multiple legal texts beginning in 2013. Observing the regulatory framework in 2017 allows us to capture full implementation of the 2010 package of reforms; and observing the framework in 2007 gives us a pre-crisis basis for comparison.

We find that the prudential regulatory framework for banks in the UK increased from almost 400,000 words to over 720,000 words. A continuity in legal drafting style however ensured that individual rules did not become more complex in terms of the concepts or operations. The network of cross-references between rules expanded, and a tighter core emerged, centered on the EU-level regulation implementing international Basel agreements. The average rule contained

more cross-references, and the chains of cross-reference became longer.

On average, the number of cross-referenced words a reader would need to process to understand the initial rule (assuming no background knowledge) grew by over 4,000%. The growth in the number of connected words outstripped the increase in the total number of words in the framework, which grew by "only" 74%. The increase was driven by cross-references in the Capital Requirements Regulation (CRR), which implemented Basel reforms in the EU.

We also contribute towards evaluating the potential for machine-readable rules. At least a third of the 2017 rules contained vague terms that require substantial interpretation, limiting the potential for machine-readable solutions. But weak correlation between vagueness and other dimensions of complexity suggests that textual measures can help identify rules that are complex for humans but require limited interpretation, and are better suited for translation into machine-readable code.

The scope of banking regulation expanded significantly after the 2008 financial crisis. The Basel III package added new international standards for a leverage ratio and the regulation of liquidity and funding risk, and reformed the definition of capital and how capital requirements are calculated. The legal implementation of Basel III in the UK and the EU also differed from Basel II given the EU's move towards a "Single Rulebook". In 2007, the EU used Directives to articulate its expectations for banking rules.[1] But Basel III was implemented through the CRR, which binds directly on banks without requiring implementing national legislation. And a new pan-European supervisor (the European Banking Authority) developed technical standards to specify details which the CRR could not cover.

We extract text from a range of documents issued by EU and UK regulators, and create a dataset that captures legal sources comprehensively, including the structure of cross-references between rules. This allows us to do a like-for-like comparison between the post- and pre-crisis frameworks.

We define cognitive costs of regulation as processing difficulties that are likely to be resolved

---

[1]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 harmonizing", 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 EU Directives, a single document—the FSA's Handbook—gave UK banks all the rules and guidance they needed to comply with.

while the reader is processing the text contained in the rule (local complexity), or that are likely to be resolved only after accessing information outside the immediate context of the rule—for instance, cross-references (global complexity). To capture these costs we employ established measures that capture the length, number of concepts and conditional operations required to process individual rules, and measures that summarize the network of cross-references across rules.

Textual interpretation is at the core of legal arguments, and necessary for rules to remain flexible (Dworkin, 1982; Black, 1997). It is also a main obstacle to the development of machine-readable rules, as interpretation requires unspecified contextual information that is not included in regulatory texts, and may not be digitized. To map out the potential for machine-readable rules, we use a vocabulary of vague expressions that rule-writers use to ensure that there is scope for discretion (for example "reasonable" and "adequate").

To validate our measures, we carry out two exercises. First, the European Banking Authority collects clarification questions on rules from banks and industry bodies. If our measures capture cognitive costs, complex rules should be more likely to get requests for clarification. We find that rules that are more connected in the network of cross-references, longer rules, and rules containing multiple concepts and operations, are more likely to receive questions. We control for the topic of the rule to address concerns about the correlation between complexity and economic importance of a rule.

Second, to understand the link between the aims of Basel reforms and our measures, we carry out a case study on rules related to the definition of capital, i.e. which instruments are admissible to meet minimum capital requirements. The aims of the reforms were to raise the quality of capital, harmonize these rules internationally, and improve transparency. This required a more detailed description of contractual features, and conditions for certain operations such as deductions. In linguistic terms, this resulted in more words, concepts, operations and cross-references—in aggregate. But we see limited change at the level of individual rules, as the increase in overall complexity was spread over a larger number of rules, connected by more cross-references.

We then conduct the same analysis for the entire set of prudential rules. The results confirm that post-reform regulation is, in aggregate, longer, contains more rules, concepts, operations and cross-references. A comparison of the distribution for the linguistic measures of individual

rules does not show significant differences between 2007 and 2017. The distributions are skewed, with complexity concentrated in the top tenth of the distributions for length, range of concepts, and operations.

Finally, to evaluate the need for interpretation we measure the distributions of concepts with vague interpretation (for example "adequate") versus precise interpretation (for example numerals following currency indicators such as "EUR"). We find that vague terms are common, despite our limited vocabulary, with about one third of the rules containing them, indicating that they require an element of human interpretation that cannot be easily translated into programming language. Instead, specific numerical values are concentrated in no more than a tenth of the rules. Vagueness and precision however have weak correlation with other measures of complexity. For example, correlation of vagueness with degree-out (the number of cross-references) is -0.11. The weak correlations between vagueness and other dimensions of complexity suggest that a subset of rules that are complex for humans, but not necessarily for machines, could be identified. We illustrate this by comparing measures of vagueness against length and degree-out.

This paper contributes to the literature on textual complexity in financial settings, and in particular banking regulation. The textual length or number of rules has been used in different contexts: studies on the costs of "red tape" (Djankov et al., 2002; Mulligan and Shleifer, 2005; Djankov, 2009; Kalmenovitz, 2019; McLaughlin and Sherouse, 2019; Coffey et al., 2020; de Lucio and Mora-Sanguinetti, 2021); studies on rational inattention, for example with respect to the effects of complex taxation (Abeler and Jäger, 2015; Benzarti, 2020; Feldman et al., 2016); textual complexity as a form of "obfuscation" in corporate reporting, for example Loughran and McDonald (2016).[2]

We present an approach to evaluate if length is a good proxy for overall textual complexity, building on NLP and network analysis methods used to analyze legal texts (Katz and Bommarito, 2014; Li et al., 2015; Fowler et al., 2007), and exploiting a novel dataset that collates texts from different legal sources, and captures the network of cross-references between rules.[3]

---

[2]On textual complexity in corporate reporting see also 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).

[3]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).

Our results indicate that measuring length may underestimate the increase in the complexity of processing individual rules, as the increase in cross-references means rule-readers may have to follow long chains of cross-references, resulting in a much larger number of "connected rules" required to fully understand a rule. We further contribute to this literature by demonstrating how directional subnetworks can be used to distinguish between centrality and contextual requirements of a rule.

Our approach lets us measure changes in cross-referencing and network structure over time. Several strands of literature suggest that more cross-referencing may be associated with higher costs. Greenwood et al. (2017) argue that regulatory complexity arises from the interdependence between rules, which generate multiple constraints distorting bank behavior. If highly cross-referenced frameworks signal sub-optimal webs of rules that are difficult to reform, as in the evolutionary approaches in Ely (2011) and Kawai et al. (2018), sophisticated incumbents may be advantaged over new entrants. However, we can't exclude that a more connected framework is also a more consistent one, with explicit (rather than implicit) links between rules that reduce information-processing costs. Consistency could reduce regulatory uncertainty (Klevorick, 1973), and navigating such an environment is easier after an initial investment in understanding the core rules, which will then be cross-referred to repeatedly. Here, our empirical results are complementary to the theoretical approach on rule complexity in Colliard and Georg (2020), who show that longer rules are not necessarily more complex if they use less specialized language and concepts, which are similar to "built-in" functions.

Finally, we provide an initial assessment of the potential for machine-readable rules. Studies of the application of machine-learning in the law tend to bypass legal language, and train algorithms to learn from patterns in the behavior of judges and other decision makers (Kleinberg et al., 2015, 2018). But this restricts the application of algorithms to standardized processes that require limited interpretation.[4] Our approach helps suggest that there may be scope to automate complex rules by translating legal language into machine-readable code, as part of a broader approach to automate regulatory processes (Branting, 2017).

The remainder of the paper is structured as follows. Sections 2 and 3 describe our measures

---

[4]Sartor (2002) argues that legal reasoning requires at least two types of background knowledge—namely, (i) of the relevant factors and (ii) of the underlying values used to assess them. To model legal reasoning, we need to formalize both, and introducing values allows comparing a larger number of scenarios.

and data, respectively. Section 4 presents our main results, comparing textual complexity before and after the Basel III reforms, and Section 5 concludes.

2   Measures of textual complexity

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 (for example Gibson, 1998). Given that we will be largely taking the perspective of rules' readers (language comprehenders) rather than that of rules' 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.[5]

We take the liberty to adopt and adapt terms from the discourse processing literature, and categorize rule 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 rule 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 rule—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.

We provide first a motivation for using textual and network measures (Figure 1) to analyze the post-crisis banking reforms, then we describe our measures in detail.

2.1   Local complexity measures

To read a rule successfully, what comprehension processing tasks does the reader have to accomplish? Upon encountering a word, the reader first has to accurately decode its orthography,

---

[5]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.

retrieve its phonological code, access its lexical entry, syntactic category, and meaning from long-term memory. Having recognized the word, the reader next needs to integrate it into the sentence currently being processed, and update their present representation of the text, namely, what has gone before. We do this incrementally as we proceed through the sentence, re-analyzing if necessary, before performing sentence wrap-up (that is updating and checking the discourse model).

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.

In the economics literature, the processing costs of complex legal texts have been expressed in terms of the number of eventualities (or states of the world) detailed in a rule (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 more primitive sentences and more logical operators. Colliard and Georg (2020) draw instead on a parallel between rules and coding algorithms, and distinguish between logical "operators" and "operands" required to describe events and behaviors. In this paper we are analyzing all banking regulations, not only those which can be operationalized as in Colliard and Georg (2020).[6]

The complexity of code is also analyzed from the perspective of processing costs. Code should be readable by others and easy to analyze to ensure it is less prone to errors. Some local

---

[6]Other, non-textual approaches exist in economics and finance literatures for assessing the complexity of statistical model-based elements of financial regulation, for example analysis of internal ratings-based models in Aikman et al. (2014). However, in this paper we focus on textual complexity.

complexity metrics used in software engineering include the length of the coding framework, the length of code in a module and cyclomatic complexity which is assessed by the number of conditions present (Somerville, 2016).

In sum, linguistics, economics and computer science have comparable approaches to textual complexity, which is represented as a function of length, number of concepts, and operations on these concepts. We capture these aspects of complexity with measures of length, lexical diversity, and conditionality.

Length   This measure attempts to capture the amount of linguistic material a reader has to retrieve, integrate, and temporarily store until sentence (or rule) wrap-up is performed. There are many ways to computationally operationalize 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 simple way to approximate this concept.

Lexical Diversity   Comprehension is facilitated when words are repeated, and 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 operationalize this concept using (1) a measure of the number of *unique words* and (2) a measure of *lexical diversity*, which measures the proportion of unique word types in a rule. 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 based on Maas' index (Maas, 1972):

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

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 do not exist in order to be able to understand them (we can process facts better than hypothetical events); and ii) if there are many different conditional clauses, we

8

have to integrate many different exceptions, which may interfere with our ability to understand the applicability of a given rule. 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).*

2.2   Global complexity measures

To process information successfully, 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 do not 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.

Measures of global complexity capture two different sources of contextual information: other rules cross-referenced in the regulatory framework; and additional (unspecified) information that is likely to be necessary to interpret vague terms contained in the rule. Our measures are based on work in network science (degree and PageRank) and the law and economics literature (vagueness and precision). Similar measures exist in software engineering such as depth of inheritance tree (DIT) and number of children (NOC) which measure the length of cross-reference chains and the number of cross-reference chains respectively (Somerville, 2016).

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 (2020) point out that rule writers can simplify the algorithm by introducing "built-in" functions, intermediary results that define concepts that are then used in other rules. Our structural network measures capture entangling between rules: the meaning of a rule with high

PageRank (or degree) affects the meaning of rules that cross-refer to it (directly, or indirectly via a chain of cross-references).

Degree and PageRank are based on rule-to-rule connections using cross-references. We summarize the regulatory framework as a network, where each rule[7] is a node and each cross-reference between rules is an edge between nodes. A rule (node) with more cross-references (edges) is more complex because a reader must leave the rule and visit other rules (nodes) to fully comprehend the initial rule.

Inward and outward networks Starting from any given set of rules ("initial nodes"), two different networks can be generated. The network based on inward expansion identifies all nodes cross-referring to the initial nodes and expands in this direction until no further cross-references are found. The network based on outward expansion identifies nodes which the initial nodes cross refer to. The network then expands until no further cross-references are found.

Take for example, the following nodes and edges: $A \rightarrow B \rightarrow \mathbf{C} \rightarrow D \rightarrow E$. If we start with Node C, the inward expanded network consists of nodes A, B and C, while the outward expanded network consists of nodes C, D and E. The implications and relevance of each network are distinct. The inward expanded network can be used to assess the centrality of a rule. Centrality measures can assess the number of rules (through chains of links) a given rule would impact if policymakers made a change to it. The outward expanded network can be used to assess the explicitly referenced context of a rule. Network complexity measures can be used to assess the number of rules a reader will have to travel to and comprehend to understand the rule they started from.

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. Incoming (in-degree) and outgoing (out-degree) edges are counted separately.

Chain length measures the average length of chains originating from a node, where a chain is an interrupt series of cross-references pointing in the same direction. For example, node A has a chain of length (outward) of two if $A \rightarrow B \rightarrow C$.

---

[7]Or each *article* in our analysis of EU texts, as described in the "data sources" section.

Connected nodes measures, for each node, the number of nodes that are connected to it via chains of cross references. For example, A's outward network has four connected nodes if $Z \leftarrow X \leftarrow A \rightarrow B \rightarrow C$. While degree counts only direct cross-references (in this case, to X and B), connected nodes also count indirect cross-references (in this case, to Z and C).

PageRank Page et al. (1999) summarizes the centrality of a node within a network. Simplifying somewhat, PageRank counts the number and quality of cross-references *to* a rule to estimate how important the rule is. More important, i.e. more central, rules are likely to get more direct and indirect links from other rules. Degree is a more intuitive measure, but only captures direct links between nodes. PageRank takes into account the whole web of indirect links that point towards a node, so more completely captures how any given node influences all other nodes in the network. PageRank also provides a ranking of nodes within the same network. As the networks in 2007 and 2017 have different size, we multiply PageRank for each node by the total number of nodes in the network in 2007 and 2017 to facilitate comparison.

Reverse PageRank Gleich (2014) summarizes the explicitly referenced context requirement of a node within a network. To simplify, Reverse PageRank counts the number and quality of cross-references from a rule to estimate how much further context is required to understand the rule in question. Rules that require more context are likely to create more direct and indirect links. It is calculated by using PageRank on the transpose of connections in the network.

2.3 Vagueness

The law and economics literature focuses on the need to interpret the meaning of vague terms contained in a legal requirement. In this literature a "rule" is a requirement that contains precise terms, while a "standard" is a requirement that achieves the same aim with vague terminology.[8] Interpretation of vague terms requires accessing additional sources of information (which are often not explicitly referenced) 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 as examples of precise terms.

Our vagueness and precision measures use hand-crafted lexicons. Requirements that use only precise terms require little additional contextual information outside of the requirement 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 requirement itself, 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 requirement. Drawing on domain expertise, we create a lexicon on 22 vague terms and compute their frequency in each requirement (normalizing for length). The lexicon includes the following vague terms: *appropriat*, adequate, available, effective, equitabl*, fair, good, likely, material, necessary, particular, possibl*, potential, practicabl*, prompt, reasonabl*, regular, several, significant, substantial, sufficient, timely.*

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

---

[8]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).

3 Regulatory texts and data

3.1 Elements of our dataset

We create a new dataset that includes the universe of prudential legal obligations that apply directly to UK-authorized banks.[9] For the purposes of this paper, we describe this as the "prudential regulatory framework". 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. It includes any mandatory requirement that binds directly on a UK-authorized banks. We limit the analysis to "banks" defined as firms subject to the EU's Capital Requirements Directive, i.e. deposit-takers including building societies (but not credit unions).

To evaluate the effect of the reforms that followed the 2008 financial crisis, we use November 2007 and November 2017 as observation dates for the pre- and post-crisis prudential regulatory frameworks.

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. Rules 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, and the Financial Conduct Authority the conduct 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".

---

[9]The code is available at https://github.com/bank-of-england/PRArulebook. It scrapes the structure, text, and network data from the PRA rulebook website.

3.2 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 rules that are in and out of scope for our analysis—for example prudential (in scope) and conduct (out of scope) rules, or banking (in scope) and insurance (out of scope) rules —it also required judgments on which rules within these source documents are relevant for our analysis.

Figure 2 summarizes the legal sources for UK banks in 2017, which capture full implementation of the 2010 package of Basel III reforms. We restrict our analysis to legally binding rules that apply *directly* to UK banks (shown by arrows). We include the EU's Capital Requirements Regulation (CRR, 2013) and associated Technical Standards (2013 onwards) made by the European Banking Authority (together, these implemented the 2010 Basel III reform package in Europe). At the UK level, we include the PRA Rulebook to capture UK rules implementing relevant EU Directives and UK own-initiative policies. We exclude the text of EU Directives, for instance the Capital Requirements Directive which formed part of the EU's legislative package (CRDIV) for implementing the 2010 Basel III reform package. But by including the UK rules which implement them, we capture the text of the legally binding rules they create for UK banks.

For 2007, we include the FSA's Handbook, which captures both the legally binding rules which implemented all relevant EU Directives in the UK, and other relevant UK rules.[10]

UK regulatory frameworks in 2007 and 2017 also included expectations for how banks should comply with binding requirements (supervisory "guidance"). Guidance was published in the FSA's Handbook in 2007 and in separate PRA "supervisory statements" in 2017. We restrict

---

[10]For 2007, identifying the comparable set of rules was more challenging as the FSA's Handbook did not clearly distinguish between: 1) prudential- and conduct-related rules; and 2) rules applying to banks versus other sectors of firms. To identify prudential rules, 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 rule in the FSA Handbook as at 27 February 2013 was prudential, conduct, or both. We also used text-matching techniques to identify rules that appeared in the FSA Handbook at both 27 February 2013 and at 16 November 2007 (our observation date for 2007 texts).

our analysis of local and global complexity to binding requirements only, i.e. our results exclude guidance. But for completeness, we include guidance when we discuss the change in the total length of the entire framework between 2007 and 2017.

The text of the prudential regulatory framework has a hierarchical structure that aggregates rules 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), paragraph 1.[11]

We define our unit of observation for local versus global complexity as a *rule*: a package of words articulating a single binding requirement. Rules reflect their writers' decisions on what constitutes the most granular level of aggregation of legal language. These units are called "paragraphs" in EU texts (CRR and associated Technical Standards) and "rules" in UK-originated texts (PRA Rulebook and FSA Handbook).

EU texts often consolidate multiple, related paragraphs into an *article*. 43% of cross-references in the CRR refer to articles. So when analysing global complexity, we use the article as our unit of observation for EU texts to avoid missing a significant part of the network structure (i.e. each node in our EU legal network is an article, comprising at least one rule). For the PRA Rulebook and FSA Handbook, we use the rule as the unit of observation when measuring both local and global complexity because the vast majority of their cross-references are between individual rules.

---

[11]Paragraph 1 reads: "Subject to Articles 93 and 94, institutions shall at all times satisfy the following own funds requirements: (a) a Common Equity Tier 1 capital ratio of 4,5 %; (b) a Tier 1 capital ratio of 6 %; (c) a total capital ratio of 8 %."

4 Results

4.1 Validation of Measures: Evidence from European Banking Authority Q&As

To what extent do our measures of linguistic complexity actually measure what they are supposed to be measuring? Public and private institutions can submit clarification questions about CRR articles to the European Banking Authority (EBA). These questions and corresponding answers are then published on the EBA website (https://eba.europa.eu/single-rule-book-qa/search). This data comes in a semi-structured form and includes information about the source of the question, topic, and the corresponding CRR article (or even paragraph in some cases). In total, 1818 Q&As were scraped from the EBA website, and we used these to validate the measures of complexity that we describe in Section 2.

The number of questions asked can be thought of as a proxy of complexity. Complexity can come from the content or the form of the article. This paper focuses predominantly on the form. Even if complexity as described in this paper is not entirely responsible for the number of questions raised about given CRR article, we would still expect that rules which are more complex (according to our measures) would lead to questions being raised about it.

In order to test this claim, we fit a logistic regression model to predict whether a question will be raised about a particular CRR article.

There were 1,450 (out of 1,818) questions related to the $n = 519$ articles CRR in our dataset. We treated the outcome variable as a binary outcome (question vs. no question). In total, there are $n_1 = 261$ (= 50.3%) articles with at least one question asked and $n_0 = 258$ (= 49.7%) articles with no question asked.

Using the above-mentioned data, we built a (ridge) logistic regression model with the following specification:

$$\text{logit}(\Pr[Y_i = 1]) = X_i + \gamma_t + \epsilon_{it},$$

where Y_i is a dummy that is equal to 1 if rule i has a Q&A attached; X_i is a vector of complexity measures for rule i; and γ_t is a vector of dummies for topic t (for example, definition of capital, liquidity, operational risk, and so on). Controlling for topic helps address concerns that questions on a given rule may be driven by the importance of the topic (for instance, in

16

terms of financial costs for banks), rather than its complexity.

Our model attained a reasonable predictive accuracy on a held-out portion of data: 62.58% (test set baseline = 50.3%). Figure 3 gives standardized coefficient estimates for the logistic regression estimated by ridge penalization (for clarity, and since we are not interested in them, we do not show topic features). The points are the coefficient estimates for a given feature and the bars extending from the points are the 95% confidence intervals around the coefficient, which we estimated via bootstrap methods. Black bars denote relevant variables, gray bars irrelevant variables.

Features relating to global linguistic processing (that is our network measures of degree out (in), PageRank, reverse PageRank) are important. Specifically, the more important or central an article, the more likely a question is posed about it. On the whole, articles point to other articles, so any queries that the former raises are typically answered by the latter. This may explain why degree-out is not itself significant. However, the more central an article is, the less likely it is to point elsewhere, and so queries that this central article raises simply go unresolved and need separate clarification.

Features around processing bulk and operation (exception) processing all result in an increased likelihood of question-posing, each with similar coefficient estimates—that is, the harder an article is to read and integrate, the more likely a question is asked about it on the EBA website.

Precision and vagueness also relate to question-posing, but perhaps not in the way we expected. Precise articles tend to elicit more question-posing, and this might be because firms have questions around specific numerical quantities. What of the negative coefficient for vagueness? Possibly, readers know that these terms are not going to be explained in any more detail—that is, they are vague on purpose to retain discretion.

Thus, based on this evidence, our measures *do* seem to be measuring what they are purported to be measuring, providing support that they are indeed valid proxies for the linguistic processing of regulatory rules.

4.2 Basel reform of rules on definition of capital

In the last section we found some evidence for the validation of our measures. Now, to explain how in practice our measures are linked to the rule-making process, we start by applying them to a subset of rules for which we can construct reasonable priors on how textual complexity ought to have changed post-crisis, given what regulators have said about pre-crisis weak spots and their objectives for post-crisis reform. This allows us to show (qualitatively) relationships between regulators' efforts to achieve specific policy aims and changes in textual complexity.

We use the subset of rules defining which instruments are eligible as regulatory capital for banks. The 2008 financial crisis revealed that the quality of bank capital had declined, as banks had raised cheaper, lower quality (Tier 2) capital to meet requirements (de Ramon et al., 2016; Herring, 2018). Credit losses and writedowns hit banks' retained earnings, but other capital instruments that had qualified as capital did not absorb losses.[12] Definitions of capital proved inconsistent across jurisdictions, and banks' disclosures were not complete enough to allow the quality of capital to be compared across banks (Basel Committee on Banking Supervision, 2020).

Accordingly, post-crisis reforms aimed to raise the quality of banks' capital resources, to harmonize eligibility criteria internationally, and to improve transparency by requiring more comprehensive and detailed disclosures (Financial Stability Institute, 2020). The Basel III package clarified the roles of different forms of capital (Tier 1 vs. Tier 2) and defined "Common Equity Tier 1" (CET1) as the highest quality of regulatory capital; introduced specific classification criteria for the different components of regulatory capital; and harmonized deductions and valuation adjustments at the CET1 level. Figure 4 shows the increase in the ratio of Tier 1 to total capital for UK banks in the years after the 2008 crisis, also as a result of Basel reforms. The change in the Tier 1 ratio is more evident for large banks, which could issue Tier 2 instruments more easily than smaller banks.

---

[12]From a regulator's perspective, capital represents banks' ability to absorb losses whilst remaining a "going concern" (Farag et al., 2013). Key characteristics that liabilities must satisfy to provide such loss-absorbing capacity are: 1) perpetual status (banks need not repay original investment from capital investors); and 2) distributions to capital investors are not obligatory, for instance, dividends to shareholders. Tier 1 capital consists of shareholders' equity and retained earnings. Tier 2 capital includes revaluation reserves, hybrid capital instruments and subordinated term debt, general loan-loss reserves, and undisclosed reserves.

We restrict our analysis to the rules defining Tier 1 capital. To make the Tier 1 definition more prescriptive and internationally harmonized, policymakers would have needed to write rules that discriminate more precisely and comprehensively between different states of the world and contractual features. Achieving this might require policymakers to introduce new concepts and use more conditional clauses to specify eligibility criteria. Less frequent use of vague terms might also be necessary to ensure consistent application of the rules. An overall increase in the length of Tier 1 rules, and higher scores for our lexical-diversity and conditionality measures in the post-crisis sample of Tier 1 rules, might be consistent with the intended effects of the reforms.

Alternatively, if policymakers tried to maintain (or simplify) the language of individual rules, we might expect increases in the number of rules relating to Tier 1 capital, and the length, conditionality and lexical diversity across rules relating to Tier 1 capital in aggregate, without changing significantly for individual rules on average. The latter approach might also generate more textual cross-references between individual rules. We might expect individual post-crisis rules to have longer chains of textual cross-references across more source documents.

We identify 17 rules defining Tier 1 capital in 2007, and 69 rules in 2017 (a 306% increase). For these rules, we start with results for local complexity, i.e. processing difficulties that are likely to be resolved while the reader is processing the text. Panel A of Table 1 shows that the number of tokens (words), unique words and conditional terms used in rules defining Tier 1 capital increased between 2007 and 2017. Total words rose by 760% (from 1,030 to 8,861), the number of unique words by 253% (from 268 to 945), and the number of conditional terms by 538% (from 13 to 83).

Panel B of Table 1 shows results for global complexity, i.e. processing difficulties that are likely to be resolved only after accessing information outside the immediate context of the rule. Starting from the rules defining Tier 1 capital (the core nodes), we construct an inward and an outward network. The inward (outward) network follows chains of cross-references pointing to (originating from) each of the core nodes (see section 2). The number of nodes in the inward network rose by 196% (from 69 to 204), and the number of edges by 153% (from 108 to 273). The increase in the size and connection of the inward network indicates that rules on definition of capital became more central to the understanding of a larger number of rules.

The number of nodes in the outward network rose by 343% (from 28 to 124), and the

19

number of edges by 433% (from 24 to 128). The increase in the outward network indicates that understanding the rules on definition of capital required following cross-references to a wider number of rules.

Vague expressions are much more common than precise ones, both in 2007 (10 and 1, respectively) and 2017 (32 and 6, respectively).

These results indicate that, for rules on definition of capital, both *aggregate* local and global complexity increased between 2007 and 2017. To understand how the complexity of *individual* rules changed, Table 2 compares means for our measures of complexity. Of the local complexity measures in Panel A, the mean number of words increases from by 74% (from 73.6 to 128 words), a substantial increase, but lower than the overall increase in the number of words in Table 1. Relative conditionality increases by only 8%, and lexical diversity decreases by 3%. The relative stability of the measures of local complexity can be explained by the increase in the number of rules. Complex material was divided across a higher number of rules, each with roughly similar complexity.

The network measures in Panel B show that the number of direct connections (captured by degree) to core nodes on definition of capital increased for both the inward and outward networks (by 278% and 212%, respectively). The chains of cross-references originating from the core nodes however shortened on average by 50% for the inward network and 19% for the outward network. As a result, the overall network complexity captured by mean PageRank and Reverse PageRank increased less than suggested the direct connectivity captured by degree.

Finally, the mean frequency of both vague and precise expressions declined (by 67% and 60%, respectively).

These results show that tightening the rules on definition of capital led to higher local and global complexity at an aggregate level. Two rule-drafting techniques limited the increase in complexity for *individual rules* on definition of capital: 1) writing more rules limited local complexity; 2) shortening the chains of cross-references limited global complexity. In the next section, we explore whether similar results also apply when we consider the whole regulatory framework.

4.3 Basel III and the UK regulatory framework

The case study in the previous section showed how specific policy aims (for rule on definition of capital) can affect textual complexity. To capture the cumulative effect of Basel III on textual complexity, we now expand our analysis to the entire prudential regulatory framework for UK banks.

4.3.1 Aggregate comparison and network maps

We estimate that, in 2017, after most of the regulatory reforms were incorporated, the universe of prudential rules and supervisory guidance for UK banks included over 720,000 words, compared to almost 400,000 words pre-crisis in 2007 (see Figure 5). Whereas in 2007 only UK-originated rules were directly binding on UK firms (including those implementing EU Directives), in 2017 the majority of rules were EU Regulations and Technical Standards that bound UK banks directly, without implementation through UK rules. The change led the volume of UK-originated rules to fall between 2007 and 2017, whilst the volume of UK-originated guidance increased.

Table 3 compares the entire prudential regulatory frameworks in 2007 and 2017. We show comparisons separately for binding rules only, and for the total of rules and guidance (we include guidance in this section's aggregate comparison for context, but exclude it in the analytical results reported in subsequent sections). Panel A shows that if we consider only binding rules the increase in length is from 222,000 words to 446,000 (over 100%). The number of conditional expressions increased from 2,250 to over 3,800 (+69%), with a similar relative change if we include guidance. The number of unique words however did not increase proportionally to the total number of words. For rules, unique words rose from over 8,700 in 2007 to almost 9,700 in 2017 (+12%, with a similar relative increase if we also consider guidance). This is not surprising, since the likelihood of repeating words increases with the volume of words.

Panel B compares the numbers of network nodes and edges. Nodes correspond to individual rules and lines of guidance and edges to cross-references between individual rules and lines of guidance. For rules, the number of nodes increases slightly from almost 2,400 to over 4,100 (+73%), with a similar change in the number of edges, from about 2,500 to almost 4,300 (+69%). This result indicates that, when we consider only rules, the network of cross-references becomes

more connected between 2007 and 2017. PRA guidance is issued in self-contained documents (Supervisory Statements), which contain fewer cross-references than pre-crisis FSA guidance, which was embedded in the Handbook, together with binding rules. If we add guidance, the number of cross-references increases only slightly (12%).[13]

The number of vague expressions contained in rules rose in line with the number of words (103%), and the number of precise expressions also increased, but not as fast (80%).

Figure 6 provides a visualization of the network for 2007. Each node corresponds to a rule in the FSA Handbook (guidance is not included), and each edge corresponds to a cross-reference. Figure 7 shows the network for 2017, including the legal source for each node: CRR (yellow nodes), Rulebook (blue), EBA TS (red) and Supervisory Statements (green). 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 or two edges. CRR Articles are at the center of the network.[14]

These results at aggregate level show that the post-crisis reforms increased the textual complexity of the prudential framework. The 2017 framework contained more words, more unique words, and more conditional words; and more individual rules and lines of guidance, with more cross-references between them.

4.3.2 Change in textual complexity of individual rules

The previous section provided evidence that the prudential regulatory framework *in aggregate* was more complex after the reforms. In this section we want to test whether *individual* rules have, on average, become more complex.

We focus on binding rules only, and exclude guidance. Guidance is non-binding, and strictly speaking it is not "law" and not enforceable as such. It is hence difficult to compare the language of guidance with actual statutory language: guidance is not drafted with the same care to principles of statutory interpretation and structure, and tends to be written in a more plain-language style, with the goal of making the statutory rules accessible.[15]

---

[13]We do not show inward and outward networks as these can be constructed only from a subset of the entire network.

[14]For visual clarity, only nodes with at least one edge are displayed. For the same reason, we do not show cross-references to entire CRR Articles (which contain several rules), leading to an under-representation of the centrality of CRR. Edges are unweighted, that is they do not show multiple cross-references between two rules.

[15]We thank Cristie Ford for suggesting to focus on rules.

We discuss first results for local and then global complexity. For local complexity, individual rules are the relevant textual unit, but for global complexity we capture CRR Articles (which contain several rules) to include cross-references that refer to a whole Article. We also compare the complexity of different legal sources in 2017 (in Section 4.3.3).

Our local complexity measures capture comprehension of individual rules *while* they are being read, abstracting from other sources of information that the reader may need to access at a later stage. Figure 8 compares the distributions of our first set of measures in 2007 versus 2017. Distributions for each year are obtained by ranking rules using the relevant measure and then plotting the mean for each decile.

The mean number of tokens (words) is similar in 2007 and 2017 for all the deciles, with the exception of the last one. In the top decile, rules are longer in 2017 (partly due to rules that include long lists and tables), but the difference is not statistically significant. The distributions for lexical diversity appear virtually identical in 2007 and 2017. Mean conditionality is higher in 2017 in the top half of the distributions (in particular in the eight and tenth deciles), while rules in the bottom half include no conditional statements.

To summarize, the distributions of local complexity measures for individual rules appear similar in 2007 and 2017, with post-reform concentrated in the top deciles of the distribution for lexical diversity. This result seems to indicate a continuity in the style of writing rules over time, suggesting that the authors try to limit the complexity of individual rules. Authors may decide to break a rule that is particularly complex into two or more different rule. In particular, the distributions for lexical diversity are similar in 2007 and 2017, suggesting that rule-writers try to limit the number of concepts in a rule.

Figure 9 presents the distributions for the measures of network complexity, for both inward and outward networks. Mean degree increased in 2017 for the top, but not the bottom deciles, resulting in more skewed distributions. Mean degree for rules in the top decile of the distribution for the inward (outward) network was 15 (18) in 2017, compared to 6 (6) in 2007. Similarly, the length of chains increases in 2017 in the top half of the distributions, with average length in the tenth decile for the inward (outward) network of 14 (13) in 2017 versus 8 (8) in 2007.

The distributions for PageRank and reverse PageRank are much more similar in 2007 and 2017, with differences only in the tenth decile. This can be explained by two considerations. First, the correlation between degree and the length of chains is weaker in 2017 (from 0.42 to

0.39 for inward links and from 0.37 to 0.13 for outward links), so that rules with high degree were less likely to have long chains. Second, both PageRank and reverse PageRank compare the relative centrality of nodes within a network, so absolute differences in terms of network connections are muted.

Figure 10 presents the distributions for the number of connected nodes. The number of connected nodes aims to capture the importance of a node in a network, however, unlike (reverse) PageRank, it is an absolute measure, and does not weigh nodes. In our context, the number of connected nodes in the inward network reflects the number of rules that cross-refer (via chains) back to the relevant rule. In 2017, on average, over 488 rules cross-referred back to the rules in the top decile, compared to 39 in 2007. For the outward network, rules in the top decile cross-referred to (and hence required reading) another 463 rules, compared to 37 in 2007.

The increase in the number of connected nodes was concentrated in the top deciles, with limited change in the bottom half of the distribution, but also affected mean values. If we consider the outward network generated starting from each rule, on average, the number of cross-referenced rules a reader would need to process to understand the initial rule (assuming no background knowledge) grew by 2,142% (from 7 in 2007 to 157 in 2017). For the inward network, the number of rules that were affected by (i.e. cross-reference to) the initial rule grew at the same rate.[16] The growth in the number of connected rules outstripped the increase in the total number of rules, which grew by "only" 73%.

Three factors help understand the large increase. First, both degree and chain length increased rapidly. Second, connected nodes can overlap, and point to (originate from) a similar set of rules. For example, 460 rules cross-refer (directly or indirectly) to Article 36 CRR ("Deductions from Common Equity Tier 1 items"), and so would require reading Article 36 to be fully processed.[17]

We can do a similar calculation for the total number of words contained in each set of connected nodes, by multiplying the total number of nodes connected to an initial rule with

---

[16]An inward connection for a node will be also be an outward connection for another node. As a result, both the total and the mean number of connected nodes is the same for both inward and outward networks. For individual rules the inward and the outward network are however different.

[17]Repeated reading is not necessary if a rule can be memorized after the first reading, but this is unlikely. In reality, several readings will be required (but perhaps not 460).

the average number of words contained in a node. The number of total words in the framework increased by 74% but, on average, the number of "connected" words that are affected by the initial rule (in inward network) grew by 4,294%. In the outward network, the number of connected words that must be read in order to fully process the average rule grew at the same rate, on average. The large increase in words is partly influenced by the change in cross-referencing style for CRR. In 2007, all cross-references (in the FSA Handbook) were rule-to-rule, while in 2017 CRR contained a large proportion (43%) of cross-references to Articles, each containing several rules, so that each such cross-reference pointed to (required reading) all rules in the same Article.

Finally, Figure 11 compares distributions in 2007 and 2017 and presents results for vagueness and precision measures. We start with vagueness, which simply picks terms from a dictionary of terms that are deliberately vague, and commonly used in regulation to ensure regulatory discretion. About a third of all rules contain at least one of our vague terms, and the 2007 and 2017 distributions are similar. Only rules in the top decile contain any of the precise terms (currency and per cent), and the proportion falls in 2017 compared to 2007.

4.3.3 Legal sources

Our data for the post-reform regulatory framework are extracted from three legal sources: CRR, EBA Technical Standards, and the PRA Rulebook. These texts perform different, complementary, legal functions, and are written following different governance processes (each of which could affect textual complexity). In this section, we concentrate exclusively on 2017 rules (in 2007 all rules were implemented in the FSA Handbook).

Figure 12 compares the distributions for measures of local complexity. Distributions for rules in CRR and EBA Technical Standards (both originated at EU level) are almost identical. The distributions for the PRA Rulebook are also similar, but contain somewhat shorter rules, with a higher level of lexical diversity.

The top two charts in Figure 13 plot degree and PageRank for inward networks of references pointing towards CRR articles and rules in the PRA Rulebook (we exclude Technical Standards, which have almost no incoming links). CRR articles in the top deciles have higher degree and PageRank than rules in the PRA Rulebook . This result supports the visual analysis in Figure 7 that the core of network of cross-references in 2017 is composed mainly of CRR requirements.

The bottom two charts in Figure 13 plot results for the outward network of references originating from CRR, PRA Rulebook and Technical Standards. CRR articles tend to have higher degree in top deciles (but the difference to Technical Standards is not statistically significant in the tenth decile), and reverse PageRank, which is also consistent with a more central role for CRR. The distributions of network measures for the PRA Rulebook tend to be below those of CRR and Technical Standards, indicating a more peripheral position in the network of cross-references.

Figure 14 compares distribution of measures of vagueness and precision across legal sources. The PRA Rulebook contains relatively more vague, and fewer precise expressions. CRR contains a relatively more precise expressions, for example key capital ratios are contained in Article 92 of CRR.

4.4   Correlation between complexity measures

The distributions of complexity measures are highly skewed, with complexity concentrated in the top deciles, and in several cases they became more skewed after the Basel III reforms. To investigate whether the distributions are driven by a limited number of rules that are highly complex on different dimensions, Table 4 summarizes the correlations between different measures in 2017. Out of 36 correlation coefficients, 15 are not statistically significant, 12 are positive and 9 are negative.

Correlations are mostly weak, and do not suggest that distributions are driven by a limited number of highly complex rules. In other words, rules that are highly complex under one dimension are unlikely to be highly complex under other dimensions. The signs of the correlations are however worth noting. Negative coefficients suggest that two dimensions of textual complexity are substitutes, and are used as alternative ways to represent underlying rule (legal) complexity. Positive coefficients indicate that two dimensions of textual complexity are complements, and are used together to represent underlying legal complexity.

We discuss briefly some of the coefficients in Table 4. The correlations between length and our other measures of local complexity, lexical diversity and relative conditionality are low, confirming that we are capturing different dimensions of complexity (our measures of lexical diversity and relative conditionality already adjust for length). The correlation of length with relative conditionality is positive (0.12), while the correlation with lexical diversity is negative

(-0.17), indicating that longer rules are associated with a higher proportion of conditional operations, but a lower diversity of concepts (more repetitions).

As expected, correlation is high (0.74) between degree-in and PageRank, two sets of network measures that capture the inward network. Similarly for the outward network, the correlation between degree-out and reverse PageRank is also high, 0.62. Correlation between degree-in and degree-out is low and negative, indicating that rules that are more central (captured by degree-in and PageRank) tend to require slightly less context (captured by degree-out and reverse PageRank).

Length is positively correlated with degree-out and reverse PageRank (0.32 and 0.18, respectively), suggesting that cross-references may be used as a way to limit the length of individual rules. Length is also correlated with degree-in and PageRank (but more weakly, 0.18 and 0.09, respectively), indicating that longer rules tend to be cross-referenced to slightly more often.

Finally, vagueness and precision are negatively correlated, suggesting that they capture alternative drafting styles, but the coefficient is small (-0.05). Neither measure is correlated with length, or other measures of local complexity. Vagueness is negatively weakly correlated with degree-out (-0.11), which seems consistent with our approach to identifying vague terms that require interpretation and judgment, rather than accessing additional information from other rules.

4.5 The potential for machine-readable rules

The correlations in Table 4 provide a starting point to think about the potential for machine-readable rules.[18] The implementation of machine-readable rules (and, more generally, of other forms of replacing human labor with machine processes, such as AI, or machine-learning) is in part a problem of identifying tasks that are complex for humans but not for machines (Athey et al., 2020).

Our measures capture different dimension of *human* cognitive complexity. For most of these dimensions, the complexity is caused by the underlying limitations of the human brain in terms

---

[18]By machine-readable rules, we mean rules that are sufficiently precise to be translated from natural language into programming language (for example by expressing the rule in terms of formal logic), without significantly changing the scope of the rule itself. See Branting (2017) for an overview of approaches to automation in the legal sector.

of language processing and memory capacity. The length of individual rules, the length of chains of cross-references, and the number of operations and concepts, all tax the human brain because they require retaining and processing information. Machines do not face the same limitations, as long the information can be digitized. One of the main obstacles is that natural language, as opposed to computer language, is often not well specified, and requires access to contextual information that might not be digitized, or even digitizable. Interpreting vague rules is cognitively costly also for humans, who may have to identify and process additional contextual information, such as precedents and market practice. However, vague terms can be valuable from a legal point of view, allowing for flexible interpretation (by adapting existing concepts, or accessing new sources of information).

In this paper, we try to capture this dimension of complexity with our measures of vagueness and precision. As mentioned in Section 2, the measures capture vague terms that rule-writers deliberately introduce to ensure supervisory flexibility (for example "adequate"), and specific numerical values that are used to define key ratios (such as minimum capital requirements) and thresholds.[19]

As shown in Figure 11 vague terms are common, despite our limited vocabulary, with about one third of the rules containing at least one vague term, suggesting that the rules in which these vague terms appear require an element of human interpretation that cannot be easily translated into machine-readable language.

If vagueness (precision) is strongly positively (negatively) correlated with other measures of complexity, this suggests that rules that are currently complex for humans will also be complex for machines. Instead, the weak correlations between vagueness and other dimensions of complexity in Table 4 suggest that a subset of rules that are complex for humans, but not necessarily for machines, could be identified. Given that our definition of vagueness is relatively strict, this approach only provides an upper bound for the potential of machine-readable rules and focuses on rules where the rule-drafter appears to have explicitly introduced vague terms in order to ensure supervisory flexibility. Human interpretation is likely to be important for the application of such rules. However, the definition of vagueness can be relaxed in future research by expanding the vocabulary of vague terms.

---

[19]Note that in reality most words have an element of vagueness, even terms that to the layperson might be considered specific, such as "assets".

To illustrate this approach, Figure 15 shows a 3D scatterplot for number of words (tokens), number of cross-references originating from a rule (degree-out), and vagueness (ratio of "vague" expressions).[20] We use number of words as a proxy for local complexity, and degree out as a proxy for network complexity, but other measures can be used. The red dots represent rules that contain no vague terms. Out of this subset, the rules that score high in terms of both length and degree-out are those that will be cognitively more complex to process for humans. For example, Article 28(1) of CRR (which defines Common Equity Tier 1 instruments) contains 747 words and 2 cross-references, but no vague terms. This rule is complex from humans to process, but does not appear to require complex human interpretation (as we have seen in Section 4.2, harmonizing the interpretation of rules on definition of capital was one of the aims of the Basel III reforms). Another example, this time from the PRA Rulebook, is rule 1.2 on Capital Buffers, which defines 14 different terms, none of which is included in our list of vague terms, and is 707 words long with 7 cross-references.

5 Conclusions

Post-crisis bank regulation reforms expanded to fix the flaws exposed by the crisis. In the EU, bank supervision was harmonized and centralized with the Single Rulebook and the creation of a Single Supervisory Mechanism. In the same period, the growth of artificial intelligence techniques has opened the prospect of machine-aided or "cyborg" supervision of banks (Proudman, 2018; Suss and Treitel, 2019).

These two developments have created new sets of issues for bank regulators. One relates to the effectiveness of post-crisis bank policy. Regulators are 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. In particular, complex regulation could create a relative advantage for large banks against new entrants, and smaller competitors.[21]

A separate, but related set of issues derives from the growing interest in the potential for

---

[20]For visual clarity, the distributions have been winsorized at the 95th percentile.

[21]But international banks may also face uncertainty about which legal framework applies in certain situations, and the relevant regulatory authority (Bassani, 2019).

machine-readable rules to increase regulatory efficiency and reduce compliance costs (Casey and Niblett, 2016). 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 more vague, more context-specific rules. 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 predictable ("tick-box") regulatory framework, with mechanical application of the rules, may be efficient in an environment with artificially intelligent agents, 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 scalable and efficient, but can also cope with the complexity of financial markets and respond to changes in the financial system and the behaviour of regulated firms (Ford, 2013).

The ultimate normative question is "How complex does bank regulation have to be to achieve regulators' objectives?" Regulators seek to control a highly complex system, with multiple connected agents, and large externalities. Because this system is adaptive—for instance, we observe herd behaviour and regulatory arbitrage—a regulatory system of individually complex and interacting parts could become progressively less effective at mitigating financial stability risks than a simpler one (Aikman et al., 2018).

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 find that that the network of rules became more interconnected, with readers having to process much longer chains of connected rules. Further work is required to understand the costs and benefits for information-processing from these changes. In particular, whether adding more explicit textual cross-references created any benefits for readers through greater regulatory consistency; and how sensitive information-processing costs are to extra cross-references.

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 (2020).

References

Abeler, J. and Jäger, S. (2015). Complex tax incentives. *American Economic Journal: Economic Policy*, 7(3):1–28.

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

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

Athey, S., Bryan, K., and Gans, J. S. (2020). The allocation of decision authority to human and artificial intelligence. *AEA Papers and Proceedings*, 110:80–84.

Bank for International Settlements (2020). Saudi G20 Presidency and BIS Innovation Hub update on the progress made in the G20 TechSprint initiative.

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

Basel Committee on Banking Supervision (2020). Basel III: A global regulatory framework for more resilient banks and banking systems.

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

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

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

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

Benzarti, Y. (2020). How taxing is tax filing? using revealed preferences to estimate compliance costs. *American Economic Journal: Economic Policy*, 12(4):38–57.

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

Branting, L. K. (2017). Data-centric and logic-based models for automated legal problem solving. *Artificial Intelligence and Law*, 25(1):5–27.

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

Casey, A. J. and Niblett, A. (2016). The death of rules and standards. *Indiana Law Journal*, 92:1401.

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

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

Coffey, B., McLaughlin, P. A., and Peretto, P. (2020). The cumulative cost of regulations. *Review of Economic Dynamics*, 38:1–21.

Colliard, J.-E. and Georg, C.-P. (2020). Measuring regulatory complexity. *CEPR Discussion Paper*, DP14377.

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

Danielsson, J., Macrae, R., and Uthemann, A. (2017). Artificial intelligence, financial risk management and systemic risk. *London School of Economics, SRC Special Paper*, 13.

de Lucio, J. and Mora-Sanguinetti, J. (2021). New dimensions of regulatory complexity and their economic cost. an analysis using text mining. *Banco de España Working Paper*, 2107.

de Ramon, S., Francis, W., and Harris, Q. (2016). Bank capital requirements and balance sheet management practices: has the relationship changed after the crisis? *Bank of England Working Paper*, 635.

de Ramon, S., Francis, W., and Milonas, K. (2017). An overview of the uk banking sector since the basel accord: insights from a new regulatory database. *Bank of England Working Paper*, 652.

Djankov, S. (2009). The regulation of entry: A survey. *The World Bank Research Observer*, 24(2):183–203.

Djankov, S., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2002). The regulation of entry. *The Quarterly Journal of Economics*, 117(1):1–37.

Dworkin, R. (1982). Law as interpretation. *Critical Inquiry*, 9(1):179–200.

Ehrlich, I. and Posner, R. A. (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.

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

Epstein, R. A. (1997). *Simple Rules for a Complex World*. Harvard University Press.

Farag, M., Harland, D., and Nixon, D. (2013). Bank capital and liquidity. *Bank of England Quarterly Bulletin*, 53(3):201–215.

Feldman, N. E., Katuščák, P., and Kawano, L. (2016). Taxpayer confusion: Evidence from the child tax credit. *American Economic Review*, 106(3):807–35.

Financial Stability Institute (2020). Definition of capital in Basel III – Executive Summary. *Bank for International Settlements*.

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

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

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

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

Gleich, D. F. (2014). PageRank beyond the Web. *arXiv*, 1407.5107.

Goldberg, L. S. and Meehl, A. (2020). Complexity in large US banks. *FRBNY Economic Policy Review*, 26(2).

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

Gutiérrez, G. and Philippon, T. (2019). The failure of free entry. Technical report, National Bureau of Economic Research.

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

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

Herring, R. J. (2018). The evolving complexity of capital regulation. *Journal of Financial Services Research*, 53(2-3):183–205.

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

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

Kalmenovitz, J. (2019). Regulatory intensity. *Available at SSRN 3451344*.

Kaplow, L. (1992). Rules versus standards: An economic analysis. *Duke Law Journal*, 42:557–629.

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

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

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics*, 133(1):237–293.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review*, 105(5):491–95.

Klevorick, A. K. (1973). The behavior of a firm subject to stochastic regulatory review. *The Bell Journal of Economics and Management Science*, 4(1):57–88.

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

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

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

Loughran, T. and McDonald, B. (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.

McLaughlin, P. A. and Sherouse, O. (2019). Regdata 2.2: a panel dataset on us federal regulations. *Public Choice*, 180(1-2):43–55.

Mulligan, C. B. and Shleifer, A. (2005). The extent of the market and the supply of regulation. *The Quarterly Journal of Economics*, 120(4):1445–1473.

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

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, pages 95–106. University of Pennsylvania Press Pennsylvania.

Sartor, G. (2002). Teleological arguments and theory-based dialectics. *Artificial Intelligence and Law*, 10(1-3):95–112.

Somerville, I. (2016). *Software Engineering*. Pearson Education Limited, 10th edition.

Sunstein, C. R. (2013). *Simpler: The future of government*. Simon and Schuster.

Suss, J. and Treitel, H. (2019). Predicting bank distress in the uk with machine learning. *Bank of England Working Paper*, 831.

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, pages 1032–1039.

TABLES

Table 1: Aggregate complexity for rules on definition of capital. The table compares the count for different dimensions of textual complexity in the sets of rules that defined Tier 1 capital (core nodes) in 2007 and 2017. The *inward* (*outward*) network is constructed following chains of cross-references *pointing to* (*originating from*) core nodes. Regulatory texts: FSA Handbook (2007), Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017).

Number of	2007	2017	Change %
Rules defining T1 capital (core nodes)	17	69	306%
Panel A: Local complexity			
Tokens (words)	1,030	8,861	760%
Unique tokens	268	945	253%
Conditional terms	13	83	538%
Panel B: Global complexity			
Inward network			
Nodes (rules)	69	204	196%
Edges (cross-references)	108	273	153%
Outward network			
Nodes (rules)	28	124	343%
Edges (cross-references)	24	128	433%
Vagueness / precision			
Vague expressions	10	32	220%
Precise expressions	1	6	500%

Table 2: Mean complexity for rules on definition of capital. The table compares the mean for different measures of complexity in the sets of rules that defined Tier 1 capital (core nodes) in 2007 and 2017. The measures are described in Section 2. The *inward* (*outward*) network is constructed following chains of cross-references *pointing to* (*originating from*) core nodes. Regulatory texts: FSA Handbook (2007), Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017).

Mean	2007	2017	Change	Change %
Panel A: Local complexity				
Number of tokens	73.6	128	54.4	74%
Lexical diversity	0.78	0.76	-0.02	-3%
Relative condionality	0.64	0.69	0.05	8%
Panel B: Global complexity				
Inward network				
PageRank	0.88	2.57	1.69	191%
Degree	1.33	5.04	29.07	278%
Length of chains	5.38	2.7	-2.68	-50%
Outward network				
Reverse PageRank	0.86	1.54	0.69	80%
Degree	0.333	1.04	8.83	212%
Length of chains	3.4	2.74	-0.66	-19%
Vagueness (x100)	1.2	0.4	-0.8	-67%
Precision (x100)	0.1	0.04	-0.06	-60%

Table 3: Aggregate complexity for the entire prudential framework The table compares the count for different dimensions of textual complexity for UK deposit-takers (banks and building societies) in 2007 and 2017. *Rules* refers to binding rules only (guidance excluded), while *Total* also includes guidance. The measures of complexity are described in Section 2. Regulatory texts: FSA Handbook (2007), Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook and Supervisory Statements (2017). Text headers and footers are not included.

	Rules				Total (rules and guidance)			
	2007	2017	Change	Change %	2007	2017	Change	Change %
Panel A: Local complexity								
Tokens (words)	221,912	445,710	223,798	100.80%	393,260	721,642	328,352	83.50%
Unique tokens	8,657	9,671	1,014	11.70%	11,799	13,420	1,621	13.70%
Conditional terms	2,250	3,802	1,552	69.00%	3,930	6,320	2,360	60.80%
Panel B: Global complexity								
Nodes (Rules)	2,376	4,106	1,730	72.80%	4,406	6,766	2,360	53.60%
Edges (Cross-references)	2,533	4,289	1,756	69.30%	4,299	4,811	512	11.60%
Vague expressions	1,371	2,785	1,414	103%	2,940	5,413	2,473	84%
Precise expressions	584	1,051	467	80%	792	1,419	627	79%

Table 4: Correlation between complexity measures. The table below shows Pearson correlation coefficients for measures of complexity, calculated at the level of individual rules (cross-references to and from CRR Articles are not included). Regulatory texts: FSA Handbook (2007), Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017).

	Length	Lex. div.	Rel. cond.	Vagueness	Precision	Degree-in	Degree-out	PageRank	Rev. PageRank
Length		-0.17***	0.12***	-0.01	0.01	0.18***	0.32***	0.09***	0.18***
Lexical diversity	-0.17***		-0.18***	0.03	-0.09***	-0.02	-0.18***	-0.01	-0.10***
Relative conditionality	0.12***	-0.18***		-0.01	-0.01	0.14***	0.06***	0.12***	0.05***
Vagueness	-0.01	0.03	-0.01		-0.05***	0.00	-0.11***	0.03	-0.03
Precision	0.01	-0.09***	-0.01	-0.05***		0.06***	-0.02	0.03	-0.05***
Degree-in	0.18***	-0.02	0.14***	0.00	0.06***		-0.05***	0.74***	-0.01
Degree-out	0.32***	-0.18***	0.06***	-0.11***	-0.02	-0.05***		-0.02	0.62***
PageRank	0.09***	-0.01	0.12***	0.03	0.03	0.74***	-0.02		0.04
Reverse PageRank	0.18***	-0.10***	0.05***	-0.03	-0.05***	-0.01	0.62***	0.04	

Figure 1: Measures of linguistic complexity
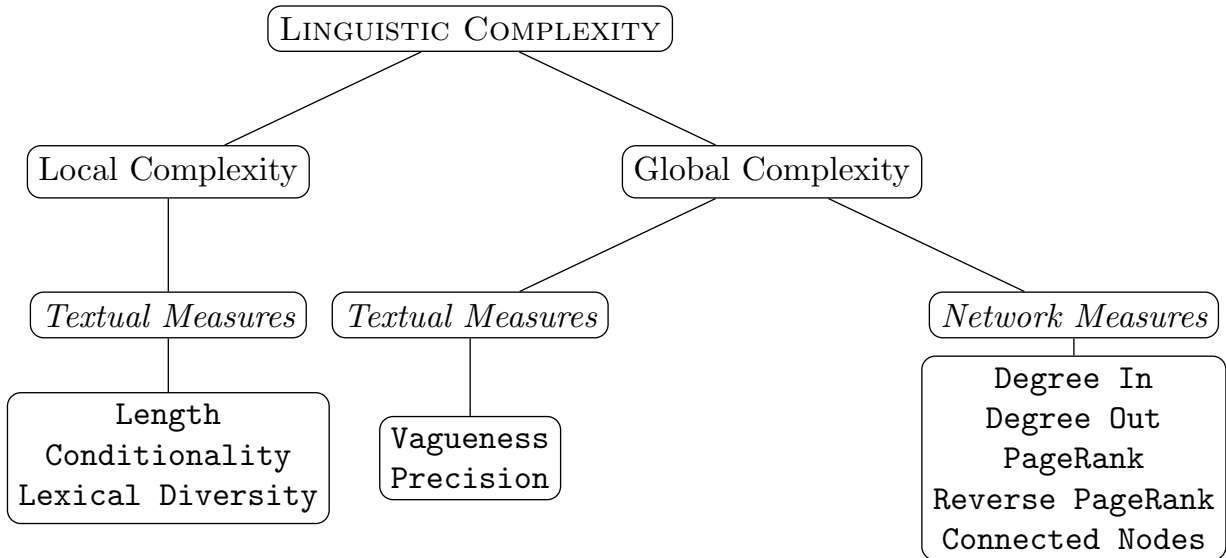
Figure 2: 2017 legal sources. The figure summarizes the legal sources for UK banks in 2017. We restrict our analysis to legally binding rules that apply *directly* to UK banks (shown by arrows): 1) the EU Capital Requirements Regulation (CRR); 2) EBA Technical Standards arising from CRR; 3) the PRA Rulebook. Supervisory guidance documents from the EBA and the PRA are directly relevant, but are not legally binding.
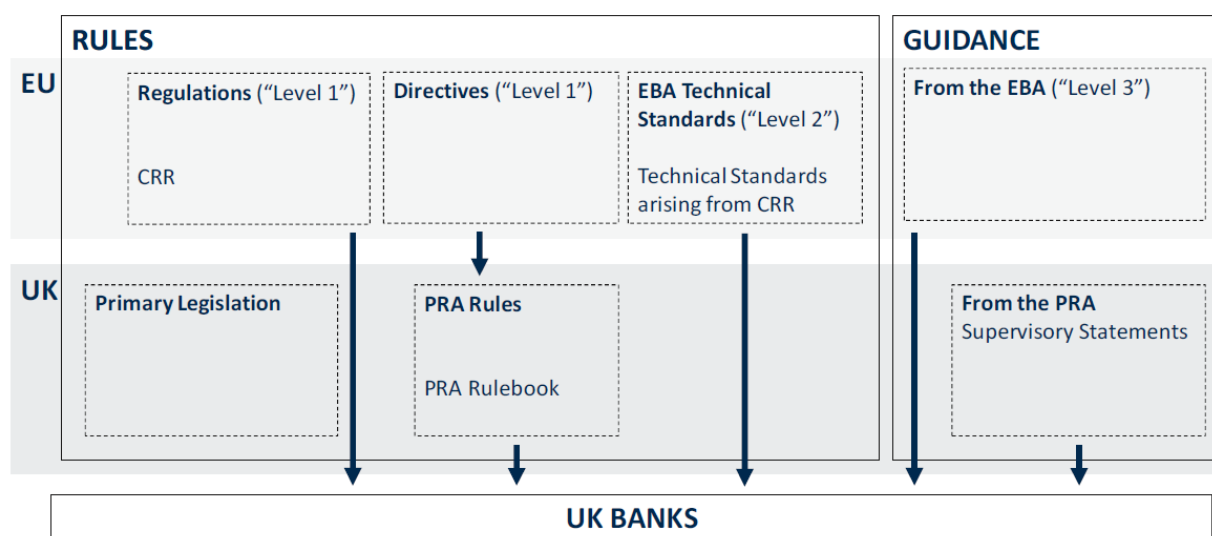
Figure 3: Scaled coefficient estimates—with 95% confidence intervals estimated by bootstrapping—for a ridge logistic regression model predicting whether an EBA article has a Q&A attached as a function of our linguistic complexity features. The model's full specification is: $\text{logit}(\Pr[Y_i = 1]) = X_i + \gamma_t + \epsilon_{it}$, where Y_i is a dummy that is equal to 1 if rule i has a Q&A attached; X_i is a vector of complexity measures for rule i; and γ_t is a vector of dummies for topic t (for example, definition of capital, liquidity, operational risk, and so on). Thus, a positive point estimate indicates that an increased value for a given feature is associated with an increased likelihood of a question being posed about an article.
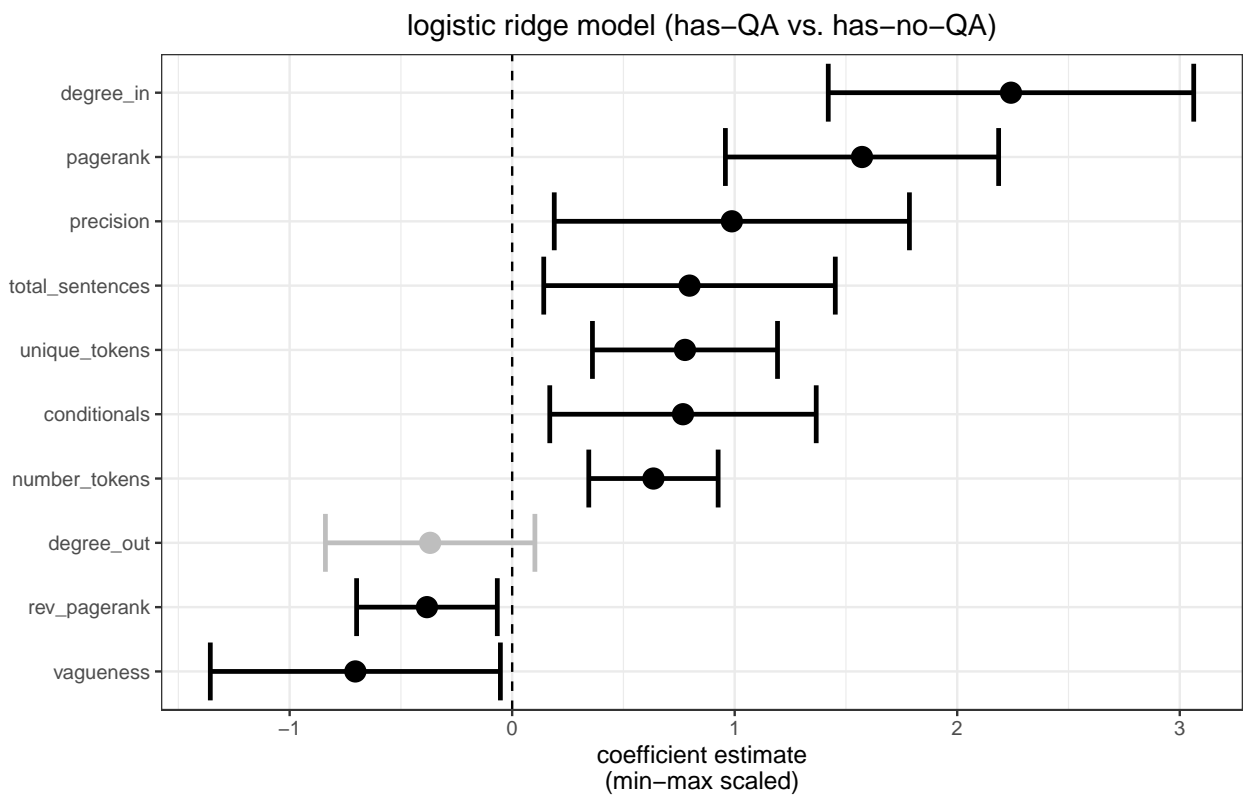
Figure 4: Tier 1 capital ratio. The ratio of Tier 1 to total capital for UK banks increased post-crisis, also as a result of Basel III reforms. The changes in the Tier 1 ratio are more evident for large banks (the six largest banking groups in the UK, excluding mutuals). Source: Bank of England Historical Banking Regulatory Database (see de Ramon et al. (2017)).
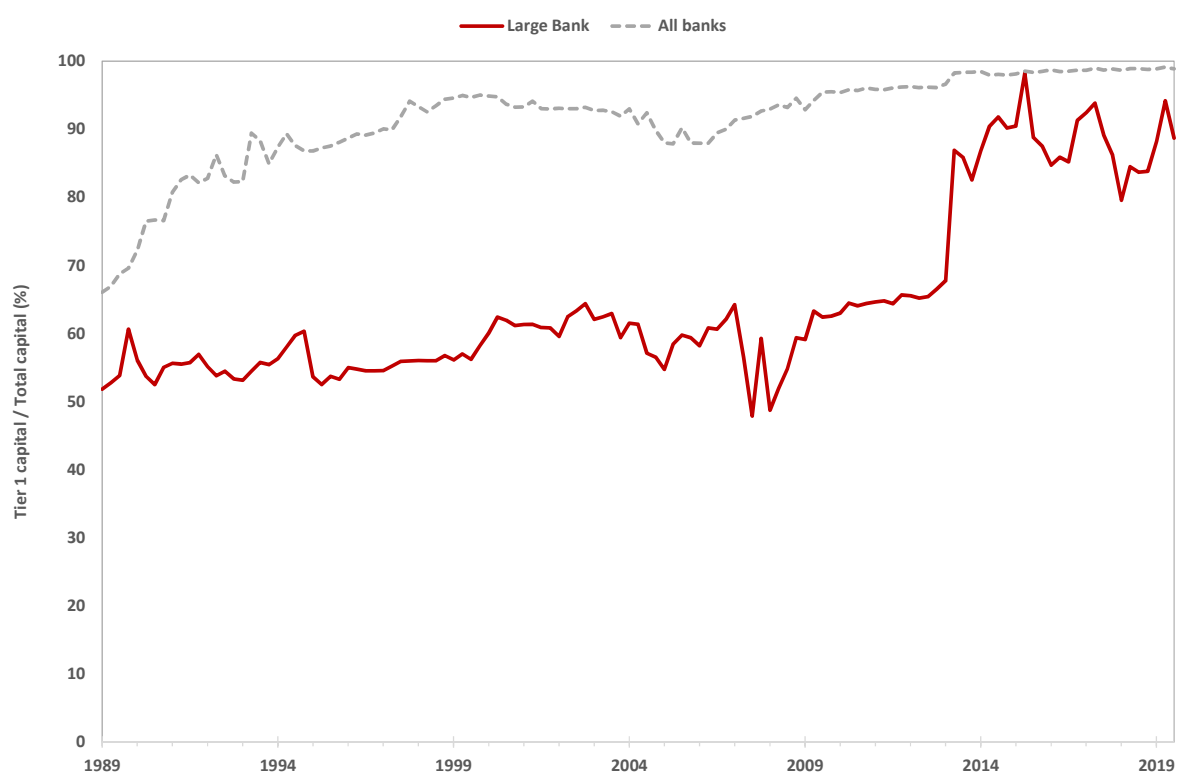
Figure 5: Length of regulatory framework. The figure compares the number of words included in the regulatory prudential framework for UK deposit-takers (banks and building societies) in 2007 and 2017. Both directly binding rules and non-binding guidance documents are shown, but not UK legislation and EU Directives that are implemented via UK rules. The legal texts for binding rules are the FSA Handbook (prudential rules only, 2007), PRA Rulebook (2017), and EU Capital Requirements Regulation and related Technical Standards (2017). There were no directly binding EU regulations or technical standards in 2007. Guidance documents are the FSA Handbook (prudential guidance only, 2007), and PRA Supervisory Statements and Policy Statements (2017).
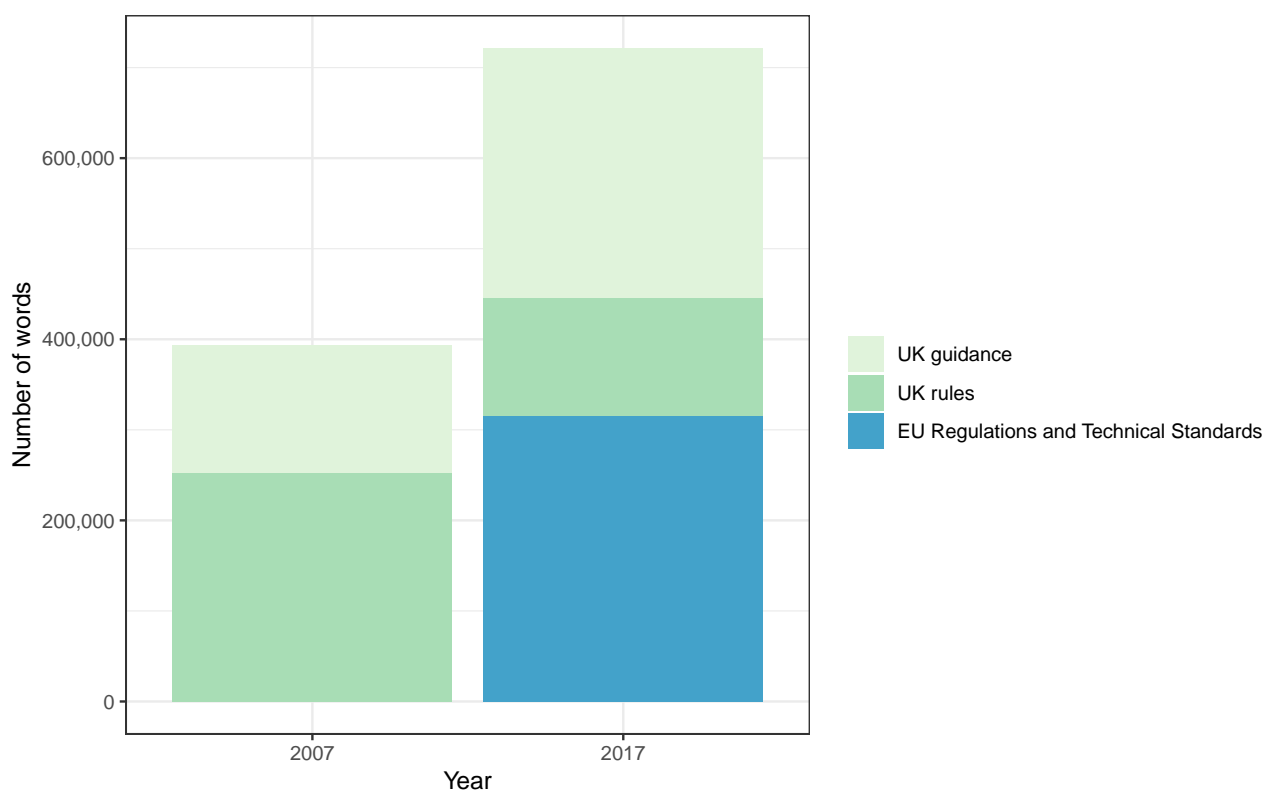
Figure 6: Network of cross-references (2007). The chart shows the network of cross-references in the regulatory prudential framework for UK deposit-takers (banks and building societies) in 2007. Each node corresponds to a rule in the FSA Handbook (guidance is not included), and each edge corresponds to a cross-reference (within and between sources). For visual clarity, only one-to-one edges and nodes with at least one edge are displayed. Edges are unweighted, namely, they do not show multiple cross-references between two rules.

Figure 7: Network of cross-references (2017). The chart shows the network of cross-references in the regulatory prudential framework for UK deposit-takers (banks and building societies) in 2017. Each node corresponds to a paragraph in the Capital Requirements Regulation (in yellow) and related Technical Standards (in red), and PRA Rulebook (in blue). Guidance documents are not included. For visual clarity, only one-to-one edges and nodes with at least one edge are displayed. Edges are unweighted, namely, they do not show multiple cross-references between two rules.
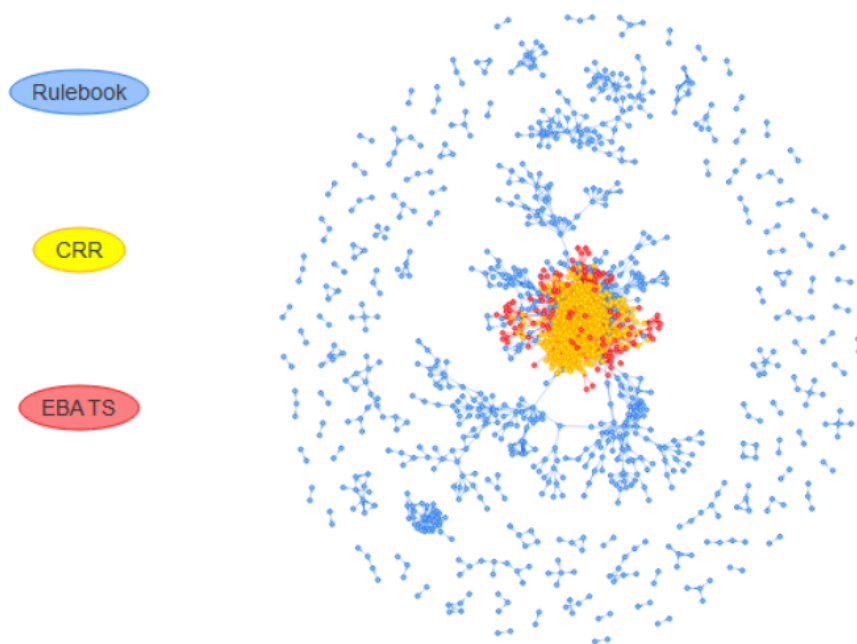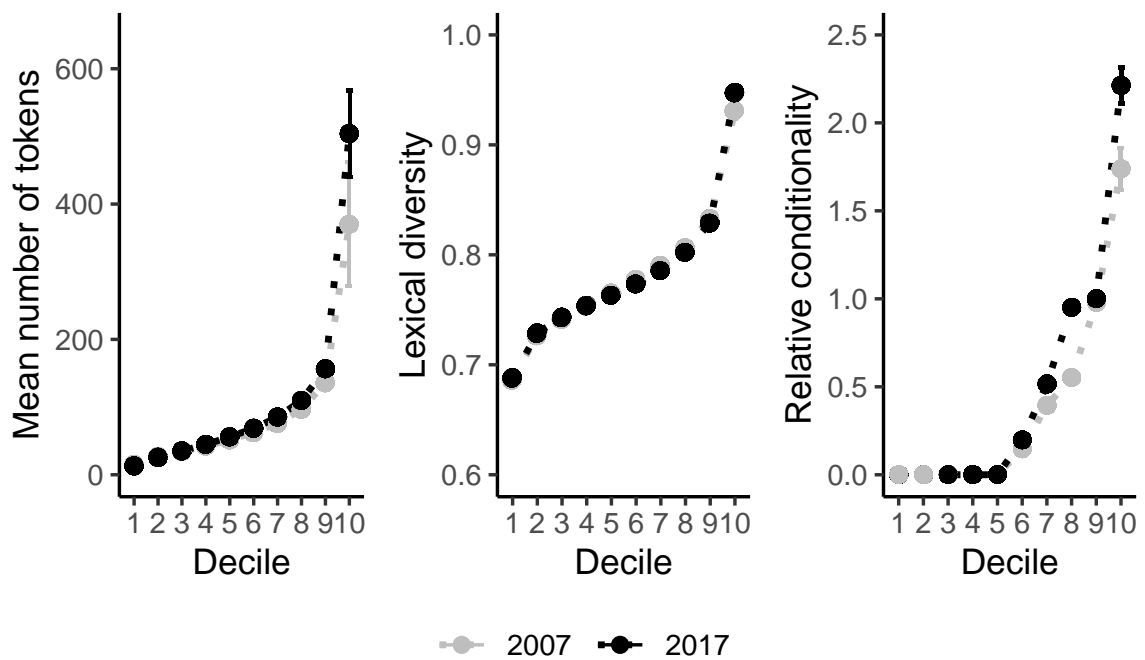
Figure 8: Distributions of local complexity measures. The charts compare the distributions of local complexity measures in 2007 and 2017. Distributions for each year are obtained by ranking rules using the relevant measure and then plotting the mean for each decile (dots) and the 95% confidence interval (vertical lines). Regulatory texts: FSA Handbook (2007), Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017). N=2,376 in 2007 and N=4,406 in 2017.
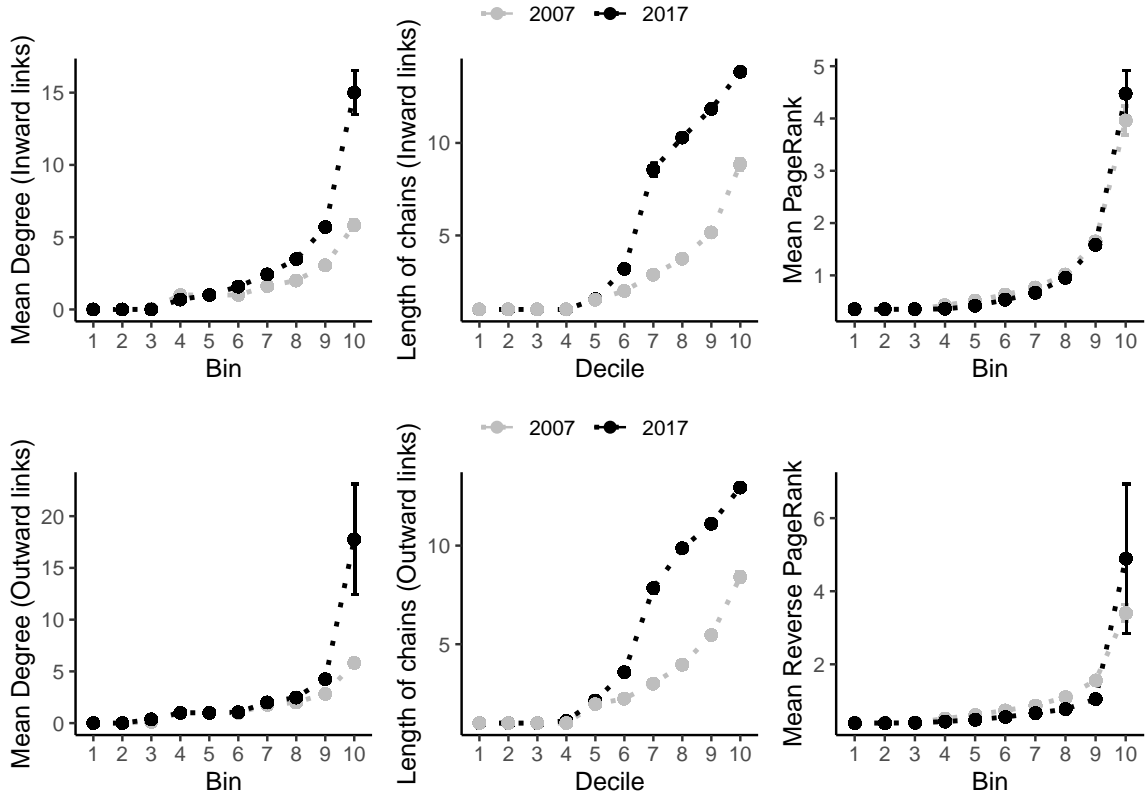
Figure 9: Distributions of network complexity measures. The charts compare the distributions of network complexity measures in 2007 and 2017. Distributions for each year are obtained by ranking rules using the relevant measure and then plotting the mean for each decile (dots) and the 95% confidence interval (vertical lines). The *inward* (*outward*) network is constructed following chains of cross-references pointing to (originating from) each node. PageRank and Reverse PageRank multiplied by number of nodes, to adjust for differences in network size between 2007 and 2017. We include both rule-to-rule edges, and edges to/from CRR Articles, which contain several rules. Regulatory texts: FSA Handbook (2007), Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017). N=2,376 in 2007 and N=2,482 in 2017.
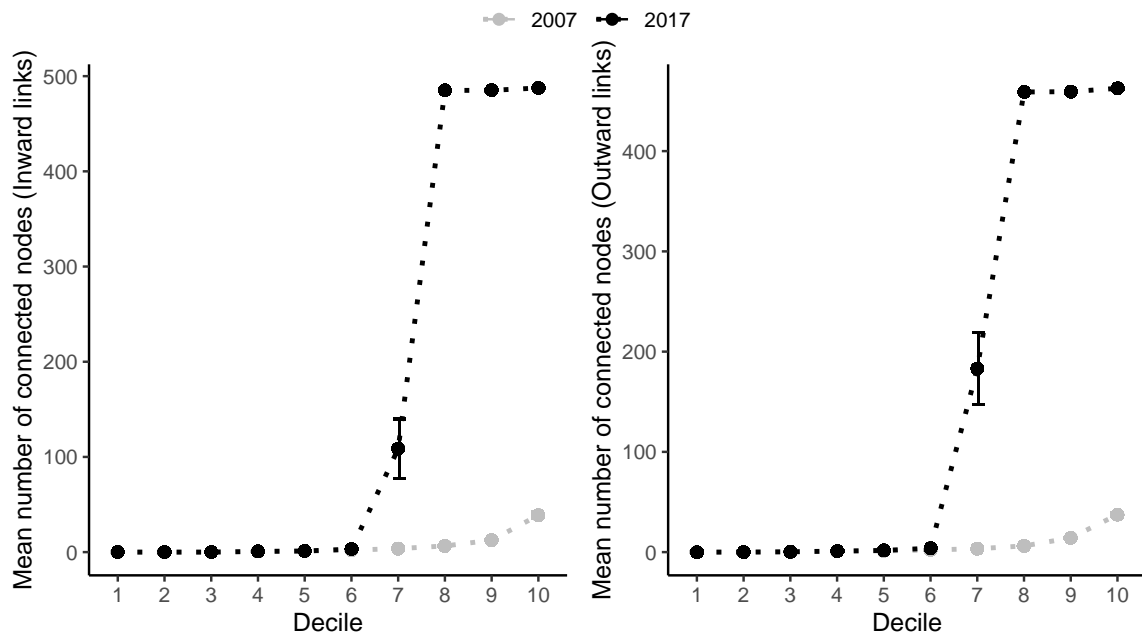
Figure 10: Distributions for number of connected nodes. The charts compare the distributions of the number of connected nodes in 2007 and 2017. Distributions for each year are obtained by ranking rules using the relevant measure and then plotting the mean for each decile (dots) and the 95% confidence interval (vertical lines).The *inward* (*outward*) network is constructed following chains of cross-references pointing to (originating from) each node. PageRank and Reverse PageRank multplied by number of nodes, to adjust for differences in network size between 2007 and 2017. We include both rule-to-rule edges, and edges to/from CRR Articles, which contain several rules. Regulatory texts: FSA Handbook (2007), Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017). N=2,376 in 2007 and N=2,482 in 2017.

Figure 11: Distributions of vagueness/precision measures. The charts compare the distributions of measures of vagueness and precision in 2007 and 2017. Distributions for each year are obtained by ranking rules using the relevant measure and then plotting the mean for each decile (dots) and the 95% confidence interval (vertical lines). Measures are based on dictionaries of vague and precise expressions. Regulatory texts: FSA Handbook (2007), Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017). N=2,376 in 2007 and N=4,406 in 2017.
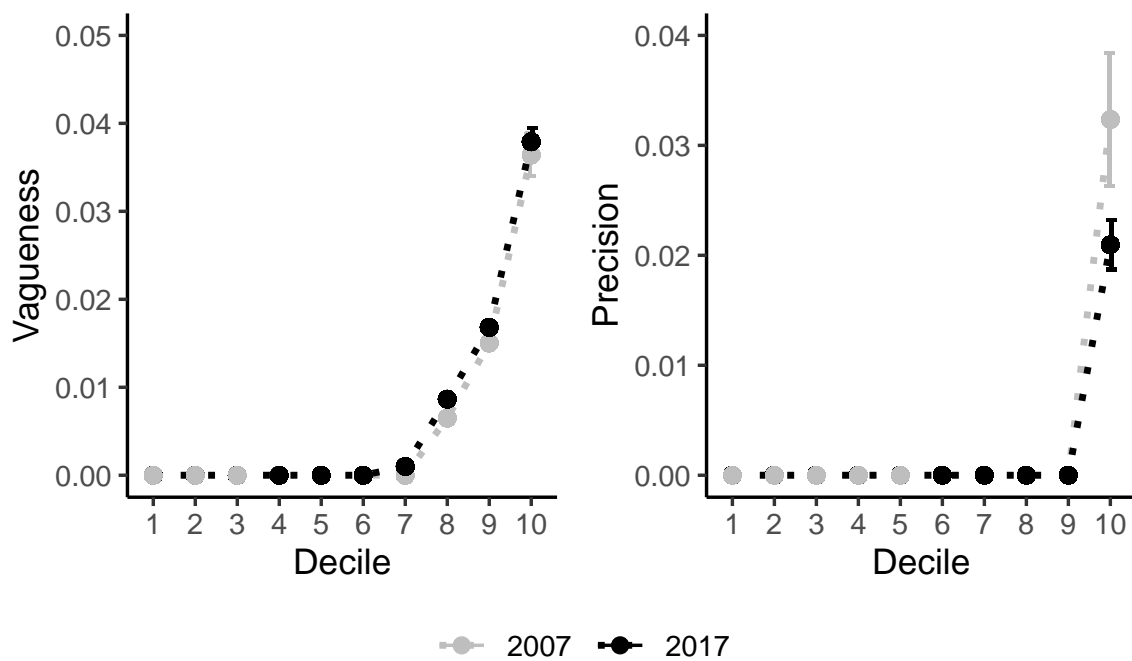
Figure 12: Local complexity measures by legal source. The charts compare the distributions of local complexity measures across legal sources in 2017. Distributions for each legal source are obtained by ranking rules using the relevant measure and then plotting the mean for each decile (dots) and the 95% confidence interval (vertical lines). Regulatory texts: Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017).
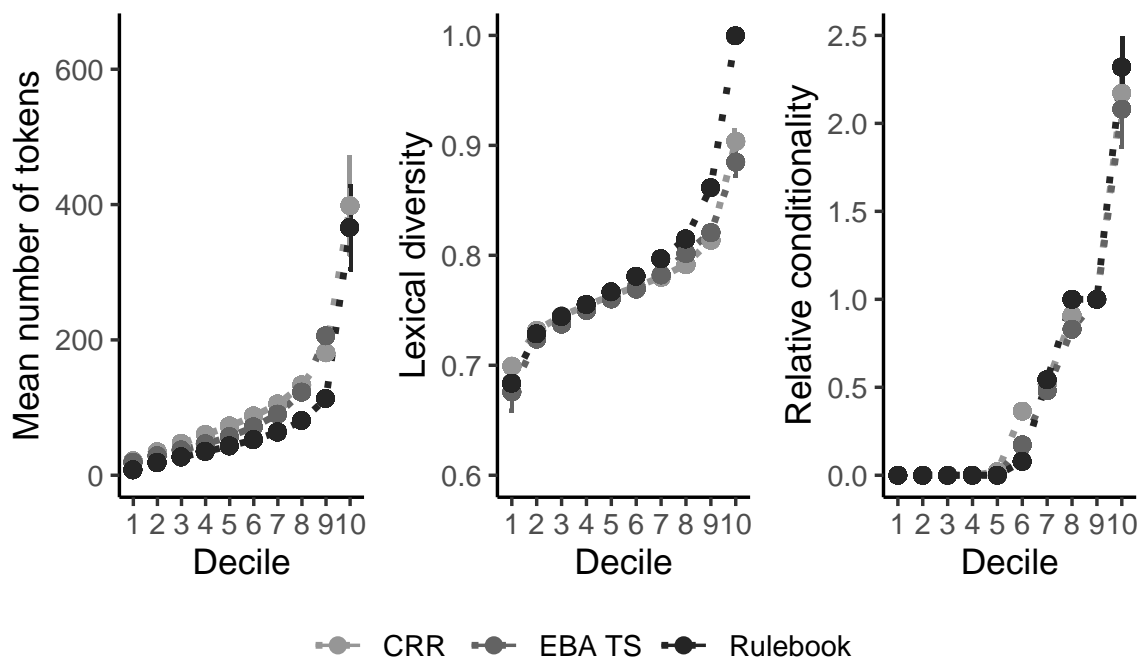
Figure 13: Network centrality by legal source. The charts compare the distributions of network complexity measures across legal sources in 2017. Distributions for each legal source are obtained by ranking rules using the relevant measure and then plotting the mean for each decile (dots) and the 95% confidence interval (vertical lines). The *inward* (*outward*) network is constructed following chains of cross-references pointing to (originating from) core nodes. Regulatory texts: Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017).
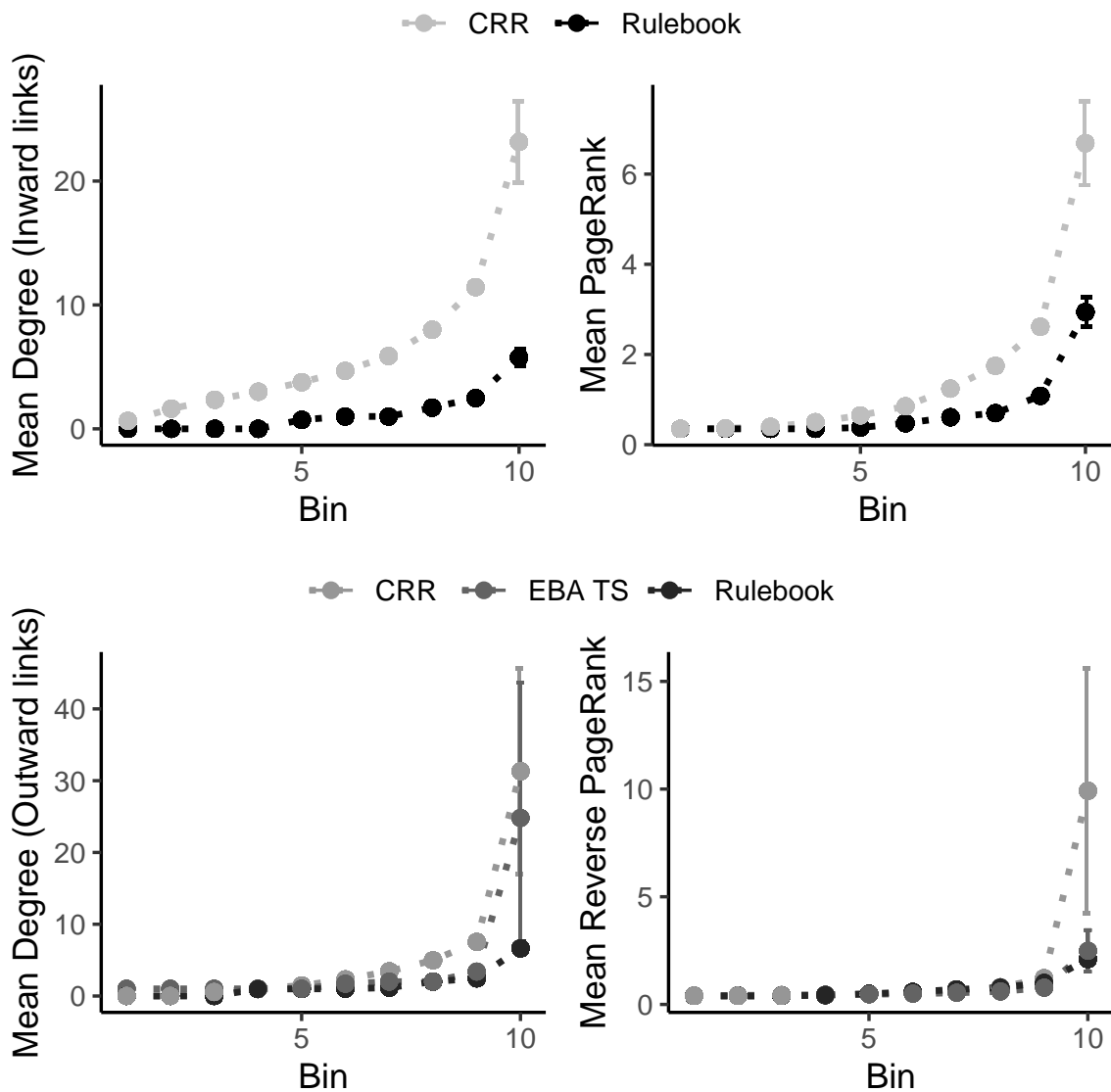
Figure 14: Vagueness/precision by legal source. The charts compare the distributions of measures of vagueness and precision across legal sources in 2017. Distributions for each legal source are obtained by ranking rules using the relevant measure and then plotting the mean for each decile (dots) and the 95% confidence interval (vertical lines). Measures are based on dictionaries of vague and precise expressions. Regulatory texts: Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017).)
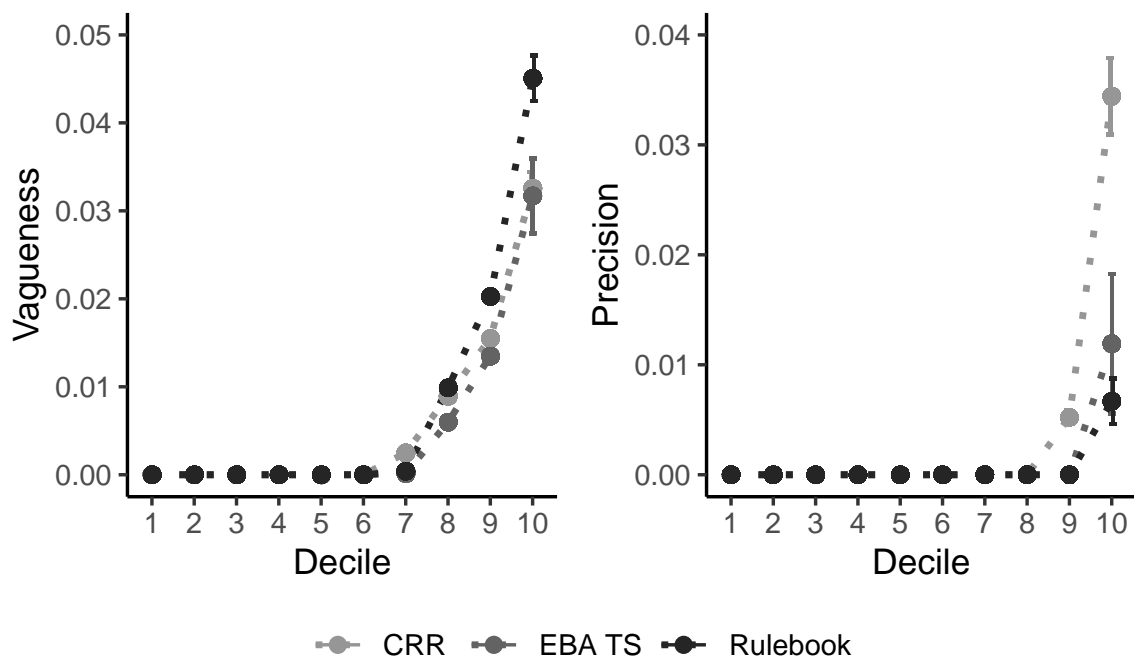
Figure 15: Vagueness versus length and degrees-out. The chart shows a 3D scatterplot for number of words (tokens), number of cross-references originating from a rule (degree-out), and vagueness (ratio of "vague" expressions). For visual clarity, the distributions have been winsorized at the 95th percentile. Regulatory texts: Capital Requirements Regulation and related Technical Standards (2017), and PRA Rulebook (2017).