

MODELING COMPLEXITY IN THE STRUCTURE OF DESIGN REGULATION

Robinson, David Myles;
Dong, Andy

Oregon State University

ABSTRACT

Government regulation shapes many aspects of the design of a product. This paper addresses the effect of the complexity of a regulation on product architecture through the structure of the regulation itself. The structure of a regulation derives from dependencies among requirements and parameters in the regulation that are ipso facto design elements. Since design elements such as requirements and parameters have no formal definition in regulation, it is difficult to identify them accurately and consistently. We apply two approaches to defining and coding requirements and parameters in the context of washing machine regulation. The two coding approaches generate networks of design elements that are analyzed to measure the complexity of the regulation and by extension the product. We find significant differences in the complexity of the regulation when coded in different ways and note deficiencies and strengths of each approach. These findings will support future research to measure the impact of regulatory complexity on product architecture.

Keywords: Design regulation, Complexity, Product architecture, Design structure matrix, Design theory

Contact:

Robinson, David Myles
Oregon State University
United States of America
robinsd3@oregonstate.edu

Cite this article: Robinson, D. M., Dong, A. (2023) 'Modeling Complexity in the Structure of Design Regulation', in *Proceedings of the International Conference on Engineering Design (ICED23)*, Bordeaux, France, 24-28 July 2023.
DOI:10.1017/pds.2023.269

1 INTRODUCTION

Historical perspectives have framed regulators and firms as a dichotomy between stifling precaution and unencumbered growth. This perspective was notably challenged in 1991 when Michael Porter proposed that well-written regulation can enhance innovative practices and competition among firms by encouraging more efficient use of resources. The Porter Hypothesis (Porter, 1991), as it is now known, suggests that environmental regulation can stimulate innovation in technology to reduce pollution and waste by de-risking investment in new technology and incentivizing companies to examine best practices for pollution control. Such policies level the playing field and constrain the design space by mandating requirement compliance. Leeuwen and Mohnen (2017) demonstrate the long-term validity of the Porter Hypothesis, but their analysis focuses on the economic factors at play, such as productivity and level of investment in new technology. While these factors remain relevant, explanations of the regulation-design relationship (including explanations for why the Porter Hypothesis holds true) lack a clear description of the specific mechanisms that constitute the relationship. When one turns to design theory literature, rich with explanations of the various factors influencing a design, the factor of regulation remains notably absent.

The European Commission—which writes the regulations used in the methods for this paper—makes no mention of the impact of regulation on design. The European Commission and Legal service (2016) provides a sort of ontology for European regulations and requests that “articles should not be too complex in structure.” However, it requests this only because complex regulations will be challenging to adjust in future amendments, ignoring any impact that complex regulation might have on design. Despite this omission, the EU publishes and enforces regulations prescribing specific requirements for countless products and systems, often describing in detail how components or parameters should interact with each other.

One well-studied concept in design theory, product architecture, has been shown to have a significant impact on design outcomes, particularly in highly complex systems. Product architecture is typically understood as the configuration of components in a product or the mapping of functions to components. To understand the direct influence of regulation on design or a product, this research begins from the premise that the content of regulations defines the functional requirements and parameters of a design. In turn, the structure and complexity of the regulation—the inter-relationships between functional requirements and design parameters—*ipso facto* defines the product architecture. The overall research goal is to study this relationship between regulation and product architecture. The first step is to quantify regulation complexity, which is this paper’s focus.

It is common to represent product architecture through the mapping of components or functions (nodes) and their dependencies (edges) in a network (Ulrich, 1995) or in its matrix-based representation, a design structure matrix (DSM) (Eppinger and Browning, 2012). Regulation could be modeled in a similar way if a framework is created with definitions for nodes and edges in the portions of regulation that relate to product architecture. However, there is uncertainty over the identification of design elements like requirements and parameters because regulations were not written with these concepts in mind. While design elements are occasionally obvious and easily identifiable, regulation does not follow a predictable form in content, structure, or writing style and there is no consistent definition for requirements and parameters.

To address this gap, this paper explores two different ways of conceptualizing the notion of requirements and parameters in regulations from a design perspective. Based upon these two perspectives, we generate two alternative design structure matrices representing the mapping between regulatory-driven requirements and parameters. We use a standard clustering algorithm to identify requirement-requirement and parameter-parameter clusters in the DSMs. These clusters are analyzed to determine the reasonableness of requirement-requirement and parameter-parameter relationships in defining a functional and product architecture, respectively.

2 BACKGROUND

This section explains the importance of design complexity and past research on regulatory complexity.

2.1 Design complexity

Product architecture and complexity go hand-in-hand. A highly interconnected architecture is more complex than one with few dependencies. While product architecture represents the manifestation of design complexity, there are innumerable external causes of dependencies, and thus complexity, in design.

While complexity typically comes with a cost, firms rarely seek to eliminate complexity—instead the goal is to manage as many sources of complexity as possible. Effectively managing complexity can mitigate undesirable affects (such as cost) while taking advantage of desirable ones (such as increased competitive advantage) (ElMaraghy et al., 2012). Complexity can be managed by modeling product architecture to identify disparate links among components and subsystems that were previously unseen. In this paper the primary modeling tool is the design structure matrix, or DSM, where system elements and interactions are represented in a rectangular matrix (Eppinger and Browning, 2012). System elements can consist of components, subsystems, or functions, but also may represent categories such as business processes or project deliverables. Elements are linked by a mark at the intersection of two elements. The flexible nature of DSMs makes them an ideal tool for comparing elements from disparate categories.

2.2 Regulatory complexity

Literature often operationalizes regulatory complexity through a text's magnitude, readability, or clarity. A policy that is initially challenging to understand might colloquially be referred to as "complicated". Endless debate in industry laments this complication. Nevertheless, public research has only recently offered specific metrics quantifying the complexity of regulation. Previous approaches to modeling regulatory complexity focus on the content of regulation instead of its structure. We have yet to find a single article addressing the impact of regulatory complexity on design or product architecture. This may be due to the nature of the problem—regulation is subject to interpretation, and modeling regulatory structure requires translation into objective categories.

Hurka and Haag (2020) proposed a set of strategies for measuring regulatory complexity based on analysis of a large set of policies from the EU. Their results show the impact these forms of complexity have on the efficiency of the policy-making process, broadly arguing that increased complexity results in increased transaction costs. While these goals differ from the goals of this research, their work defines three ways in which a policy can be "complex": structural, which refers to the size of the text; linguistic, which refers to the caliber of the language; and relational, which refers to the inter-dependencies among policies. None looks at the structure within a single regulation, and the dependent variable in their analysis—time spend writing policy—does not consider the policy's impact.

Colliard and Georg (2022) take an approach to measure regulatory complexity that is more aligned with this research, drawing on computer science to algorithmically model regulation based on "operators" and "operands". They conceptualize regulation as a set of rules or commands (operators) acting on a regulated entity such as a bank (operands). Represented as an algorithm, the regulation takes inputs from the operand and operates on it to produce a requirement. Conceptualizing the regulation this way allows the authors to apply mathematical models derived from the Halstead Measures—typically found in computer science. Much like the framework we establish in this paper, Colliard and Georg represent regulatory elements in a way that can be analyzed with techniques external to the field of policy-making.

Both studies acknowledge the importance of regulatory complexity and suggest that any effort to measure it must look beyond the world of political science. However, their approaches need to address the underlying structure of the regulation or dependencies among regulatory elements.

3 METHODOLOGY

The methodology comprises two steps to model design elements in washing machine regulation. In the first step, we develop and test two approaches for coding design elements. In the second step, we construct a DSM with the design elements gathered for each coding method to visualize the data and prepare for analysis. We analyze the DSMs to produce graphs of modularity and node degree dependency, generate metrics of perceived regulatory complexity and compare the two coding approaches.

3.1 Data set

The data for this paper come from the *Official Journal of the European Union* (EU) 2019, which publishes the official record of all EU legislation. The Official Journal contains a set of legislation known as “ecodesign requirements for energy-related products”—a group of regulations targeting consumer appliances to reduce energy consumption. The methods are conducted on “Annex II - Ecodesign requirements” from Commission Regulation (EU) 2019/2023, establishing ecodesign requirements for household washing machines and household washer-dryers. This regulation has a wide variety of design requirements and therefore met the needs for the methodological development component of this paper. Other regulations were reviewed throughout the content-coding process to develop flexible and widely-applicable rule sets. However, a complete analysis is only described for the washing machine regulation.

3.2 Modeling regulation

This section of the methodology describes the rationale and rule sets for content-coding approaches used to classify design elements and their relationships within design regulation.

3.2.1 Content coding

A structured method for identifying design elements in the text must accurately recognize design elements, be repeatable across different raters, and work flexibly with a wide variety of regulations. Content analysis fits this need as it allows for a structured description of qualitative data in varied contexts (Stemler, 2015). The academic literature provides a starting point for two robust approaches: qualitative content analysis and grammatical content analysis. The first set of rules is inspired by the conventional qualitative content analysis approach described in Hsieh and Shannon (2005), in which data is categorized through inductive analysis of the text. The second approach comes from linguistic content analysis, in which the text is analyzed through a lens of its linguistic elements. Due to the relatively consistent grammatical structures found in regulations, a rule set can be created that associates certain grammatical structures with design elements.

Each approach follows the same general process to identify design elements: the rater first reads the section of text in its entirety. Then, the rater applies rules to classify the section as a requirement or not. If the section is a requirement, the rater applies rules to identify parameters. Both approaches automatically determine dependencies if a parameter appears inside a requirement.

Approach 1, Qualitative Content Analysis:

1. Sections in this approach are analogous to sections of the regulation. Read the entire section and note its purpose.
2. Determine whether the section mandates that the subject have certain functionality. If the section does not mandate functionality, this section is not a requirement. Skip to the next section.
3. If the section is a requirement, list all parameters in the section related to the subject.

Approach 2, Grammatical Content Analysis:

1. Sections in this approach are defined as the text around a single imperative verb. Read the entire section and note its purpose.
2. Identify and mark the imperative verb phrase(s). If the section does not contain an imperative verb phrase, this section is not a requirement. Skip to the next section.

3. If the section is a requirement, mark all noun phrases and nested noun phrases in the section. Nested noun phrases qualify if they are found independent of a greater noun phrase elsewhere in the regulation.

A second rater validated the results. The second rater was trained with a portion of the results from the first rater before independently coding a new section. The two raters compared results to determine if there were any inconsistencies. Inter-rater reliability was computed to measure the repeatability of each approach. Krippendorff's alpha (Krippendorff, 2004; Hayes and Krippendorff, 2007) implemented with igraph (Csárdi and Nepusz, 2006) was chosen as the statistical measure of inter-rater reliability because it is effective even with incomplete data or partial agreement among raters.

The following demonstrates how each approach would code a sample of regulation. For the qualitative approach, parameters are listed after the section:

(1) for household washing machines and the washing cycle of household washer-dryers, the weighted water consumption (WW, in litres/cycle) for the eco 40-60 programme shall be:

$$WW \leq 2,25 \times c + 30$$

where c is the rated capacity of the household washing machine or the rated washing capacity of the household washer-dryer for the eco 40-60 programme;

The section mandates certain functionality in the design, and thus is a requirement. The parameters are household washing machines, washing cycle of household washer-dryers, weighted water consumption (WW), eco 40-60 programme, rated capacity of the household washing machine, and rated washing capacity of the household washer-dryer.

For the grammatical approach, the imperative verb phrase is marked with *italic* and parameters are marked with **bold**:

- (1) **household washer-dryers** *shall provide* a **complete cycle for cotton laundry**, named ‘**wash and dry**’:
— which is continuous if the **household washer-dryer** provides a **continuous cycle**;
— where the **washing cycle** is an **eco 40-60 programme as defined in point 1**; and
— where the **drying cycle** achieves **cupboard dry status**;

3.2.2 Design structure matrix

The design elements identified through content coding populate DSMs which provide a visual representation of the network structure of the regulation. For individual regulation, a separate DSM is created using each content coding approach. The DSMs have dimensions n by m , where n is the number of requirements and m is the number of parameters. In the body of the matrix, an “X” denotes a relationship between a requirement and a parameter. DSMs are typically constructed as square matrices representing dependencies within a single domain, such as connections among components in a system or relationships among persons in an organization. Two or more matrices from different domains can be combined to produce a multi-domain matrix (MDM). The DSMs created in this paper represent only the intersecting matrix in an MDM. The DSM's for requirements x requirements and parameters x parameters are excluded.

The network of the regulation—as defined in the DSMs—is analyzed using the R programming language to visualize and quantify regulatory complexity. The code identifies communities of nodes that are densely interconnected with few external connections (clusters) using a community detection algorithm known as the the Louvain method, then plots these as graphs. More complex networks will show fewer clusters with higher overlap. The node degree is calculated for each node to produce a cumulative distribution of node degree. The intent is to determine the homogeneity of the node degree. Non-homogeneous networks have a few nodes that are very important whereas in homogeneous networks nodes are equally important.

Req's	Parameters																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	
1.1	X	X	X	X	X	X	X																					
1.1a	X	X	X																									
1.1b	X	X		X																								
1.2	X	X	X																									
1.3	X	X	X		X	X	X																					
2.1		X	X					X	X	X																		
2.2		X						X																				
2.3		X						X			X										X							
2.4		X	X					X	X	X	X										X							
2.5		X								X																		
3.1	X	X	X					X			X																	
3.2		X						X				X																
3.3	X	X	X					X			X		X															
3.4		X						X				X									X							
4.1	X	X	X					X					X	X	X	X	X	X	X	X								
4.2	X	X	X					X					X								X							
4.3		X						X													X	X	X					
4.4		X						X													X		X					
4.5	X	X	X					X						X	X	X	X	X	X	X					X			
4.6		X						X													X	X				X		
5	X	X	X					X						X							X					X	X	
5.1	X	X												X							X							X
5.2	X	X												X	X	X	X	X	X									X

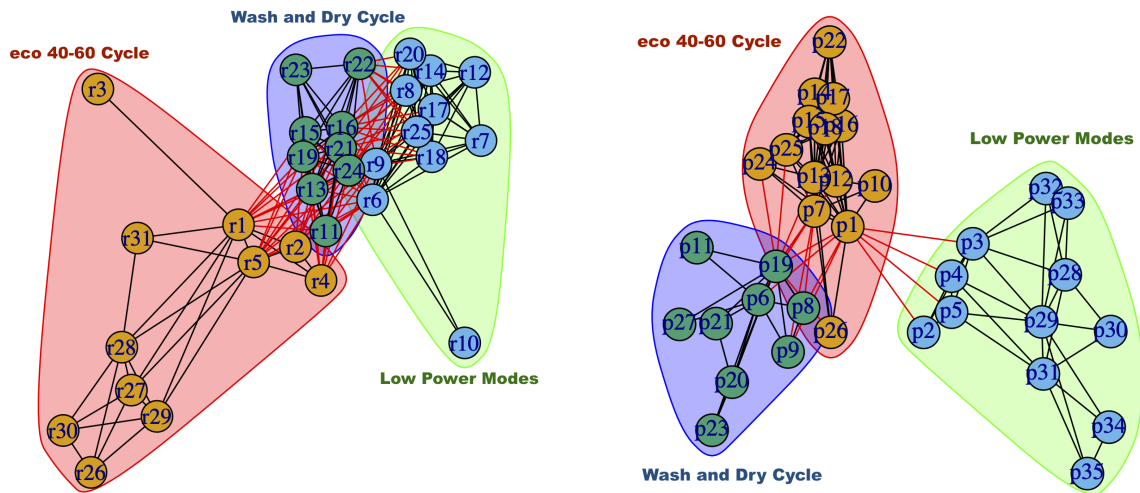
Figure 1. Portion of DSM generated from qualitative content coding method

4 RESULTS

Manual coding of the washing machine regulation produced two matrices, representing each of the qualitative and grammatical content coding approaches. The qualitative approach produced a matrix with 31 requirements, 35 parameters, and 66 dependencies. The grammatical approach produced a matrix with 39 requirements, 104 parameters, and 143 dependencies. Agreement between two raters for both approaches, measured with Krippendorff’s alpha, is 1, indicating perfect agreement among raters for n=20 samples. While both methods produced a high level of reliability, it is suspected that the grammatical approach would remain more repeatable in extended testing due to the reduced decision-making load for the rater. Additionally, the grammatical approach is structured to be analogous to how a computer program would parse the text grammatically, which would guarantee perfect repeatability. On the other hand, the qualitative approach requires more intervention during training to ensure the rater understands how to read the context of the regulation.

Figure 1 shows a portion of the DSM that was generated by the qualitative approach. The image demonstrates The parameters for ‘washing machines’ and ‘washer-dryers’ are included in the visual but excluded from the analysis in R as they are only included in the DSM for coding purposes. In general, overarching parameters such as these should be excluded from analysis because they would link all requirements and parameters, obscuring useful insight. The DSM visually shows the presence of several “bus” parameters, i.e., parameters that cross many requirements. Two of these parameters are machine cycles defined in the regulation that must be included in every machine. Many sections of the regulation impose requirements on these parameters. Another example of bus components is parameters relating to washer/washer-dryer capacity. Parameters like cycles and capacity are key drivers of complexity in this regulation.

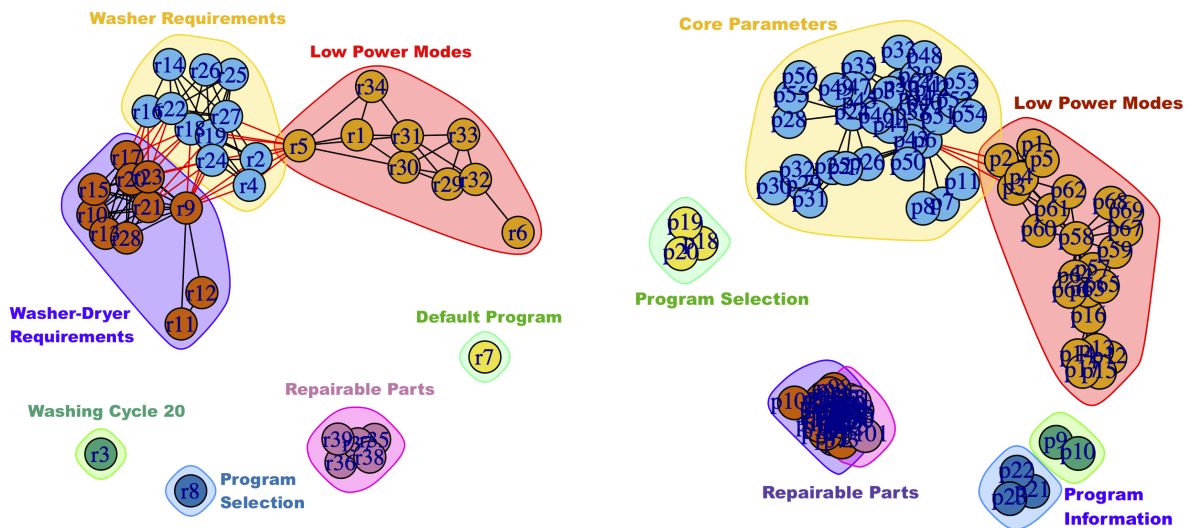
Qualitative



(a) Qualitative Method, Modularity of Requirements

(b) Qualitative Method, Modularity of Parameters

Grammatical



(c) Grammatical Method, Modularity of Requirements

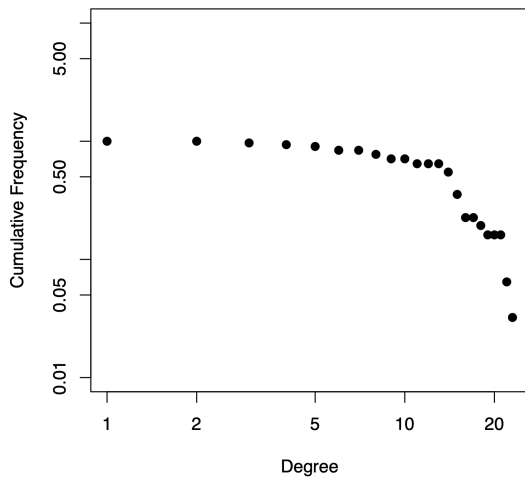
(d) Grammatical Method, Modularity of Parameters

Figure 2. Modularity analysis for qualitative and grammatical approaches

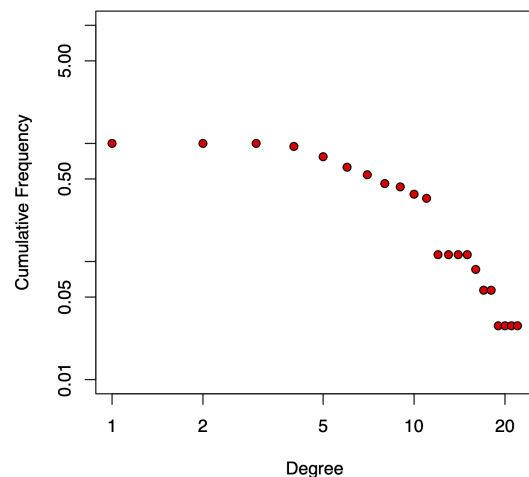
Further analysis of the network structure of the regulation reveals the complexity and allows for a comparison of the two approaches. Requirements and parameters are divided into separate graph representations to isolate complexity within each category. Two representations of complexity are presented: network modularity and node degree distribution.

The modularity of the requirements and parameters is an effective way to understand complexity by comparing the number of clusters and the overlap between clusters. Systems that have overlapping clusters are more complex than systems with distinctly separated clusters. Requirements and parameters are shown clustered into groups representing clusters in the network in Figure 2. The figure shows two graphs each for the qualitative and grammatical approaches. clusters of nodes are surrounded by a

Qualitative

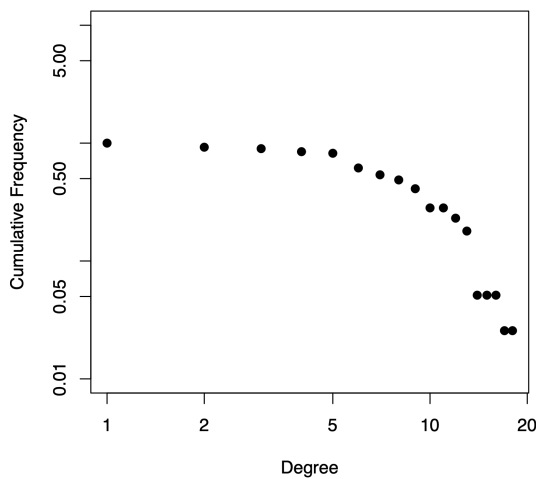


(a) Qualitative Method, Degree Dependency of Requirements

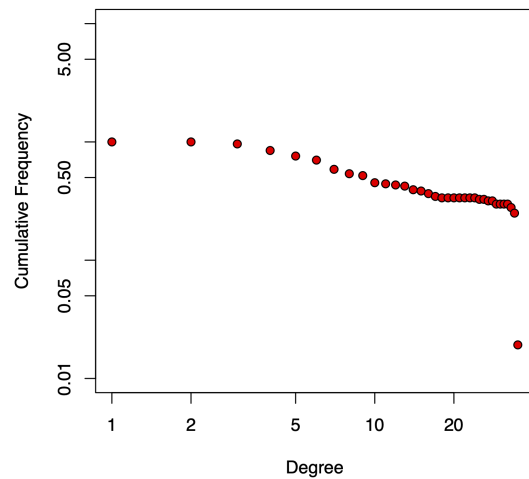


(b) Qualitative Method, Degree Dependency of Parameters

Grammatical



(c) Grammatical Method, Degree Dependency of Requirements



(d) Grammatical Method, Degree Dependency of Parameters

Figure 3. Node degree distribution for qualitative and grammatical approaches

bubble of arbitrary color. The figure makes clear the modularity and interconnections of the parameters for each approach. The qualitative approach produced graphs with fewer clusters with more overlap, indicating lower modularity and higher complexity. The graphs from the grammatical approach show six or seven clusters each, some of which are completely independent of other clusters. Additionally, the clusters with overlap do so far less than the comparable connections in the graphs from the qualitative approach. This demonstrates higher modularity and lower complexity of the regulation as determined by the grammatical approach when compared to networks generated by the qualitative approach.

Node degree in undirected networks represents the number of connections a given node has with other nodes. The number of connections is known as node degree. The node degree distribution plot shows the probability that a node in the network has a given degree. In Figure 3, node degree distribution is seen for the networks of requirements and parameters from each of the coding methods. In plots 3a, 3b and 3c, the distribution shows a high likelihood of finding nodes in the network with low degree

(fewer connections) and a slightly lower chance of finding nodes in the network with high degree (many connections) as the probability drops off. This indicates that the networks are non-homogeneous, in which many nodes have few connections and a few nodes act as “hub” nodes with connections to many other nodes. These networks are more complex. In plot 3d there is almost an equally high likelihood of finding nodes with low degree as there is finding nodes with high degree, indicating that the parameter network generated from the grammatical method is far more homogeneous than the other networks. This network thus has the lowest complexity.

5 DISCUSSION

The two content coding methods produced very different measures of regulatory complexity, indicating that the method by which requirements and parameters are interpreted in regulation leads to divergent results. Both approaches show strengths and weaknesses in terms of their repeatability and accuracy. While the grammatical method implies a simpler network structure, further examination of the underlying reasons reveals that this only occurs due to the method’s inaccuracy.

We evaluate the methods by comparing how each coding approach treats several example parameters. The first example is “rated capacity”, one of the regulation’s most influential categories of parameters. Section 4.1 in the regulation reads:

(1) for household washing machines with a rated capacity higher than 3 kg and for the washing cycle of household washer-dryers with a rated capacity higher than 3 kg, the Washing Efficiency Index (Iw) of the eco 40-60 programme shall be greater than 1,03 for each of the following loading sizes: rated washing capacity, half of the rated washing capacity and a quarter of the rated washing capacity; ([The European Commission, 2019](#))

In this section, both coding approaches pick up “rated washing capacity”, but only the qualitative approach is capable of identifying that this is referring to two types of rated washing capacity and, therefore, two separate parameters: one for household washing machines and another for the washing cycle of household washer dryers. The qualitative approach allows greater consideration of the context of the parameter, while the grammatical approach ignores context. This example demonstrates errors in both methods. In the qualitative approach, the context must be interpreted subjectively and is therefore not guaranteed to be interpreted the same way across raters. The grammatical approach consistently identified parameters, but the regulation’s phrasing obscures the parameter’s correct interpretation.

The error-prone nature of the grammatical approach inaccurately perceives modularity. The method defines *parameters* as noun phrases – this will net some meaningless or vague parameters such as “other information” and “the term ‘eco’” that may have no real connection to the design. Some phrases have more meaning: “load” is a parameter, but without context, it is meaningless. These phrases are often found in a single instance in the regulation, thus becoming independent clusters of nodes and giving the appearance of modularity.

The qualitative approach is challenging to carry out, and the level of subjectivity involved during coding further increases the time required to code a single regulation. During the coding, raters must make subjective decisions about what qualifies a requirement as relating to the design or whether two parameters written slightly differently in two places are distinct or the same. This makes large-scale data collection exceptionally difficult. An obvious progression is to write a set of rules that combines elements of both approaches to negate some of their deficiencies. For instance, the rater could first apply the grammatical rules to identify all noun phrases, then use a qualitative approach to filter out irrelevant parameters. However, the rater must still consider the context of the parameters and make subjective judgments about what qualifies as a parameter. Additionally, this hybrid approach is no easier to automate than the qualitative approach.

6 CONCLUSIONS

This paper describes the first part of a search for a clear description of the mechanisms by which regulation affects design. To do this, we compare two methods for identifying requirements and parameters

in the regulation to map the network structure of the regulation. The choice of approach makes a difference; the grammatical approach yields noise and clusters that are not as easily understandable as in the qualitative approach, indicating that any attempt to automate the classification of requirements and parameters in regulation should avoid grammatical approaches that seek specific linguistic indicators. The qualitative approach shows promise as an effective, labor-intensive process for accurately modeling the structure and complexity of regulation.

7 FUTURE WORK

Future work should include scaling of validation of the methods with complete testing on multiple regulations and further inter-rater reliability testing. In the next stage of this research, the product architecture defined by the regulation will be compared to the existing product architecture of a washing machine. The difference in complexity between the two is a new external source of design complexity known as realized regulatory complexity. This will help further evaluate the meaning and significance of the regulation-derived complexity.

REFERENCES

- Colliard, J.E. and Georg, C.P. (2022), “Measuring regulatory complexity”, *HEC Paris Research Paper No. FIN-2020-1358*, <http://doi.org/10.2139/ssrn.3523824>.
- Csárdi, G. and Nepusz, T. (2006), “The igraph software package for complex network research”, *InterJournal Complex Systems*, Vol. 1695.
- ElMaraghy, W., ElMaraghy, H., Tomiyama, T. and Monostori, L. (2012), “Complexity in engineering design and manufacturing”, *CIRP annals*, Vol. 61 No. 2, pp. 793–814.
- Eppinger, S. and Browning, T. (2012), *Design Structure Matrix Methods and Applications*, Engineering systems, MIT Press.
- European Commission and Legal service (2016), *Joint practical guide of the European Parliament, the Council and the Commission for persons involved in the drafting of European Union legislation*, Publications Office, <http://doi.org/10.2880/5575>.
- Hayes, A.F. and Krippendorff, K. (2007), “Answering the call for a standard reliability measure for coding data”, *Communication methods and measures*, Vol. 1 No. 1, pp. 77–89.
- Hsieh, H.F. and Shannon, S.E. (2005), “Three Approaches to Qualitative Content Analysis”, *Qualitative Health Research*, Vol. 15 No. 9, pp. 1277–1288, <http://doi.org/10.1177/1049732305276687>.
- Hurka, S. and Haag, M. (2020), “Policy complexity and legislative duration in the European Union”, *European Union Politics*, Vol. 21 No. 1, pp. 87–108, <http://doi.org/10.1177/1465116519859431>.
- Krippendorff, K. (2004), “Reliability in content analysis”, *Human communication research*, Vol. 30 No. 3, pp. 411–433.
- Leeuwen, G.V. and Mohnen, P. (2017), “Revisiting the porter hypothesis: an empirical analysis of green innovation for the netherlands”, *Economics of Innovation and New Technology*, Vol. 26 No. 1-2, pp. 63–77, <http://doi.org/10.1080/10438599.2016.1202521>.
- Porter, M. (1991), “America’s green strategy”, *Scientific American*, Vol. 264, p. 168.
- Stemler, S.E. (2015), “Content Analysis”, in: R.A. Scott and S.M. Kosslyn (Editors), *Emerging Trends in the Social and Behavioral Sciences*, Wiley, 1 edition, pp. 1–14, <http://doi.org/10.1002/9781118900772.etrds0053>.
- The European Commission (2019), “Commission Regulation (EU) 2019/2023 laying down ecodesign requirements for household washing machines and household washer-dryers”, *Official Journal of the European Union*, Vol. 315, p. 285–312. OJ L 315, 5.12.2019, p. 285–312.
- Ulrich, K. (1995), “The role of product architecture in the manufacturing firm”, *Research policy*, Vol. 24 No. 3, pp. 419–440.