How to apply an innovative method to analyze a policy study exploring zipf's law on korean economic policy direction reports

Eunmi Lee, *College of General Education, Kookmin University, Seoul 02707, Republic of Korea,*
foodbo@kookmin.ac.kr

Abstract. Purpose/Research Question: This study aims to quantitatively and schematically confirm the stability of the policy due to institutional restrictions and determine how the variability of the policy itself dissolves through the deconstruction and analysis of policy documents. This is based on the premise that both the institution and the policy are projected in the policy document, and these characteristics will be described later. In short, this study aims to explore the possibility of a harmonious and integrated interpretation of institutional research and policy changes by objectively describing how institutions and policies are linked and constrained.

Key Literature Reviews: Whether policy change results from political negotiations or limited choices, a debate has continued since the 1970s regarding the importance of the processes leading to these decisions, decision-making structures, social norms, ideas, and interpretations. March and Olsen [1,2] criticized pluralism, behaviorism, and rational choice in their approach to policy-making through micro-exchange relationships between individuals or interest groups. Policies are appropriately selected for social norms and conventions when there are institutional constraints in the context of formal and informal rules. Hall [3], Thelen, and Steinmo [4] asserted the rules, compliance procedures, and practices to structure relationships between policy actors in the political and economic sectors, such as institutions. The institutions influence the interests and preferences of policy actors and play a role in structuring the power relationships. Specifically, because the institutions constrain the policy-making process and political decisions, the policy is an output under institutional constraints and a path-dependent adjustment to past choices [5,6,7,8,9,10].

Design/Methodology/Approach: To demonstrate the policy changes under institutional constraints through the Zipf distribution [11] of policy documents, this study deconstructed and analyzed the 25-year economic policy direction report from 1993 to 2017.The economic policy direction announced by the government at the beginning of each year is a combination of policies and economic policy goals to be pursued during the year and which present major policy tasks. The dual structure of the Zipf distribution reflects the historical context, and the institutional situation creates policy stability and variability simultaneously. High-frequency words, subject to strong institutional constraints, are used repeatedly and in conjunction to form the basic trajectory and framework of policy documents and lead to path-dependent stability. Rare-frequency words, which are discretionary choices by policy actors who are under loose constraints, create policy variability, but such changes will be limited and temporary. Moreover, a long-term change occurs due to word substitution in the overlapping region where rare and high-frequency words intersect, but the scope of the change is narrow and gentle.

Findings/Results: In the analytical results, the economic policy direction showed a variability that changes more than a certain level each time, but it represented a modest change in the long run. The dual structure of the word frequency of the Zipf distribution is intended to have both the variability and stability of policy changes. The rare-frequency words used in the economic policy direction produced temporary and local changes in the short term, allowing the economic policy direction to adapt flexibly to the policy environment and the political situation. Rare-frequency words, under loose constraints, are easily adopted by policy actors without causing major backlash because they do not directly impede social perceptions and dominant interests in the economic policy direction.

Research limitations/Implications: First, there should be a shared consensus among researchers as to whether policy changes can be understood through word changes in the policy documents .This is a question underlying text analysis, which will require continued support from the academic community and additional methodological support. Second, this study analyzed the degree of change through the frequency of words based on the Zipf distribution, but supplementary and additional case analysis is needed to validate this reasoning. Researchers should also discuss whether the Zipf distribution, presented in terms of experience, can be directly applied to generalized discussions about policy change.

Keywords: Policy change, institutional research, stability, variability, text analysis, Zipf distribution, economic policy, duality.

INTRODUCTION

Policy actors are constantly exposed to external pressures to expect new things and self-competition to change the policies they produce each time. Policy changes may occur as a result of the intrinsic reaction of policy actors to internal and external demands for change; however, policy changes are described as part of a somewhat hopeful, bureaucratic expression. Whatever it is, institutional research sees it as a constraint on policy change [3,4,9,12].It is viewed as a change in policy within an institutional framework and structure. The system imposes restrictions on the government's ability to legislate and implement policies and influences policy choice strategies for political and economic actors. Furthermore, the institution defines the interests and preferences of policy actors and provides a framework for forming power relations with other actors. Using Ikenberry's [5] expression, policy is just a micro-adaptation process under the shadow of the past. The policy changes every time, but that change is ultimately within the institutional limits.

Accordingly, policy changes and institutions are related, but both have different views and approaches. Policy research did not easily consider institutional constraints because it analyzed relatively narrow areas of the system within a period of fewer than 10 years, and institutional research was insufficient to consider the mid-level social phenomena over a long period. Therefore, micro policy changes were overlooked. There was also a tendency for researchers to neglect each other [18].Hall [16], who studied the process of changing British economic policy from Keynesian policy to monetarist policy, merely described Kingdon's policy agenda as a footnote. Kingdon [19] did not also comment on new institutionalism, despite the development of March and Olsen's model[18]. Furthermore, Atkinson [20] pointed out that Lindblom's [21,22] incremental policy change does not consider the institutional context.

Institutional research has adopted a method to describe the specificity of the context descriptively and procedurally, depending on the researcher's insight and trained expertise [23,24,25,26].This research method prefers contextual interpretations of specific historical processes and outcomes and to perform synchronic or diachronic comparisons. This reflects the characteristics of new institutionalism, which criticizes behaviorism and pluralism and derives research results from individual behavior [1].It is inherently distanced from microscopic and quantitative approaches.

However, it is necessary to deviate from the mutual alienation of each other and to explore the intersection of policy research and institutional research. Finding an alternative to the narrative method is necessary to alleviate the habituation of the research method selection and seek to prove the connection between the system and policy. This study endeavors to demonstrate how the system constrains and suppresses policies. The aim is to quantitatively and schematically confirm the stability of the policy, due to the institutional restrictions, and to determine how the variability of the policy itself dissolves through deconstruction and analysis of policy documents. This is based on the premise that both the institution and the policy are projected in the policy document, and these characteristics will be described later. In short, this study aims to explore the possibility of harmonious and integrated interpretation of institutional research and policy changes by objectively describing how institutions and policies are linked and constrained.

RELEVANT LITERATURE REVIEW

Before describing how institutional constraints appear in policy change, the theoretical discussion is summarized as follows. Researchers in the 1950s and 1960s, the early days of policy research, understood policy as an interaction between members of society or interest groups—a vector sum of influences with different forces and directions[9]. Policies were formed through the balance between interest groups that collectively reflect personal preferences, and policy was changed as the influence between these interest groups changed. From pluralistic perspectives, the state does not have its own policy preferences but rather acts as neutral arbiters to coordinate and mediate interactions between interest groups. The institution in pluralism is only an incidental phenomenon that reflects the preferences of individuals or interest groups, and pluralism assumes actors and policies without institution.

Policy changes were approached in terms of the limitation of cognitive ability and lack of information necessary for a policy decision, rather than the macroscopic and structural constraints surrounding the policy. Lindblom [21,22,27] applied Simon's [28] study of bounded rationality to policy research; rationality-based policy-making was not only practically feasible but also normatively invalid. Therefore, incremental changes by successive limited comparisons were desirable. It was argued that all policy measures could not be evaluated due to policymakers' limited cognitive ability and lack of time and

information and because the policy was decided within a range of limited analysis and mutual satisfaction, focusing on realistically agreeable policy alternatives. Lindblom [21] maintained policy change focuses on microscopic factors and is far from macroscopic constraints.

Whether policy change results from political negotiations or limited choices, the debate on the importance of the processes and processes leading to these decisions, decision-making structures, social norms, ideas, and interpretations has emerged since the late 1970s.March and Olsen [1,2] criticized pluralism, behaviorism, and rational choice to approach policy-making through micro-exchange relationships between individuals or interest groups. A policy is considered to be appropriately selected for social norms and conventions, with institutional constraints in the context of formal and informal rules. Hall [3], Thelen and Steinmo [4] asserted the rules, compliance procedures, and practices to structure relationships between policy actors in the political and economic sectors, such as institutions. The institutions influence the definition of interests and preferences of policy actors and play a role in structuring the power relationships. Specifically, because the institution constrains the policy-making process and political decisions, the policy is an output under institutional constraints and a path-dependent adjustment to past choices [5,6,7,8,9,10].

As for the scope of institutional constraints, Hall [32] presents three levels of concepts. First, the most inclusive institution is the highest social structure, including the capitalist economic system, which includes the privatization of the means of production and the democratic political order. This is the basic framework that is the most important foundation of policy direction, and it is very difficult for any policy actor to select a policy that goes beyond these top limits. However, the mid-range system is related to the structure and framework of the social management of each country, and the mid-range system includes the election system, political system, bureaucracy, labor relations, and vocational training system. This is the basis for the policymaking differences between countries with different characteristics, depending on the context and historical background. Moreover, the narrow institution is related to practices, regulations, and procedures in public organizations, which are less restrictive than the above-mentioned institutional concept, but that does not mean that they are completely isolated from the privileged power and status of a particular group.

Yet, Hall [16], who studied the process to change economic policy from Keynesian policy to monetary policy, divides the relationship between policy change and institutions into three types. Policy changes were classified into three types: (1) first-order changes in which the outcome changes every year without changing the policy objectives and policy outputs, such as adjusting the government budget;(2) second-order changes, which are changes to the policy outputs that do not change the policy objectives; and (3) third-order changes of a paradigm change, in which the policy environment, policy goals, and policy outputs change rapidly. The first change gradually changes under institutional constraints, and in the third change, there is a punctuated change in which the stabilizer paradigm shifts to another paradigm [16].In other words, Hall's [16] first change in the policy and institutional relationship is the discretionary level change of policy actors in a limited scope, and the second change is the selection of new policy measures within the scope of policy objectives. The third changes present as changes to the concept, perception, idea, and thought frame of society as a whole.

Accordingly, from the standpoint that the policy is under the shadow of the institution, the policy has limited variability within the institutional constraints, and simultaneously, it has stability that does not rapidly deviate from the institutional framework. Of course, this discussion is not intended to determine the one-way effect on the policy of the institution. It is also the independent and dependent variables of policy changes that influence the institution and change it through policy actors' strategies and choices. It is expected that discussions on the interdependence of institutions and policies will be discussed in depth in other studies. This study focuses on approaching objectively identifying the visibility of change, one of the difficulties faced in policy research, under institutional constraints. The emphasis is on demonstrating how the variability of policy changes and the stability of institutional constraints are harmonized and realized.

METHODOLOGY AND DATA

The policy is a micro-adaptation under the shadow of the past of institutional constraints [5]; the policy has short-term variability in response to the external environment and internal demands while maintaining stability within the institutional framework in the long term. In other words, the policy seems to change every time, but when viewed from a distance, it has a dual character that appears to be stable. Depending on at what point the researcher views the policy, the policy is also variable and shows stability that does not deviate.

It is not easy to simultaneously consider the duality of this policy change. As pointed out in the introduction, due to the unique research characteristics of policy and institutional research, there is a limit to understanding each other in a converged manner. By approaching government budgets and government documents as detailed targets for analysis, errors may occur in interpreting variability and stability. Government documents, especially policy documents, are texts that present relevant facts and countermeasures to accurately recognize and make decisions about policy-related situations and problems. It is a linguistic tool that communicates and records the preferences and will of the government in an institutional context.

However, researchers are unfamiliar with the nature of policy documents and sometimes do not know. Rather, after demonstrating the policy theory for the government budget, sometimes the explanatory power will be used naturally in policy documents. It is insufficient to consider that policy documents and government budgets have different characteristics. The government budget is arithmetically displayed, so a comparative analysis is relatively easy, and statistical analysis is possible by assuming the degree of increase and decrease as a normal distribution. However, the policy document tends to maintain the overall quantity at a certain level, and the text of the policy document does not exist as a normal distribution [44]. Specifically, the government budget is additionally and limitedly changed, but the policy document changes through the replacement of the existing word with a new one. A new methodological research approach that is distinct from the government budget is necessary for policy documents composed of text.

There are three ways to understand policy documents: (1) qualitative content analysis that focuses on the meaning and structure and makes inferences;(2) discourse analysis that explores linguistic perception, norm, and interpretation from a social constructive perspective; and (3) text analysis on word patterns and statistical descriptions. To understand the duality of policy changes, this study focuses on approaches to search for the characteristics and frequency of words contained in policy documents, and a summary of this study's discussions is included below.

The first discussion is that policy documents include complex words, and changes occur as word composition and frequency change. Particularly, it is possible to measure the change in the policy document by deconstructing the policy document into words that are the smallest unit of meaning and by looking at how the words change—thereby inferring the policy change. In terms of change, the words contained in policy documents are differentially duplicated, imported, and deleted. When a certain word is duplicated repeatedly, the policy document has continuity with the past, whereas the other word gradually decreases—existing as a testimony to the past and then disappears. Furthermore, new words that have not been adopted through imitation and learning, experimental attempts, and creative thinking are introduced to replace existing units. This differential growth of words leads to variations in word frequency, which in turn leads to changes in policy documents. The words that are preferred more or less depend on the political and economic circumstances. Certain words are restricted in use in an institutional context, and words suitable for ideas [45,46,47,48,49] are politically adopted by policy actors. Also, to maintain the interests of policy actors, there is a behavior of intentionally adopting something different from words produced in the past.

The second argument is that policy documents can be expressed in vector space models of word frequency [56]. Specifically, a policy document can be converted into an algebraic form of a vector, and a specific word is represented by a non-zero number in the corresponding dimension in the vector. When the policy document $D_1, D_2, \cdots D_m$ is as shown in Equation 1 below and there are n index words in the whole, the entire word set of the policy document is expressed as an m*n matrix. This is based on the statistical semantics hypothesis that if the vector of word frequencies is similar, the meanings that people want to say are similar [56].

$$D_1 = (w_{11}, \ w_{21}, \ w_{31}, \ w_{41}, \ \cdots\cdots\cdots \ w_{n1})$$

$$D_2 = (w_{12}, \ w_{22}, \ w_{32}, \ w_{42}, \ \cdots\cdots\cdots \ w_{n2})$$

$$\cdots\cdots\cdots\cdots$$

$$D_m = (w_{1m}, \ w_{2m}, \ w_{3m}, \ w_{4m}, \cdots\cdots \ w_{nm}) \tag{1}$$

The third argument is that the words in the policy document are not distributed in a normal distribution.

The words in the policy document exist not as a normal distribution distributed evenly around the average value, but as a power distribution of the power function.
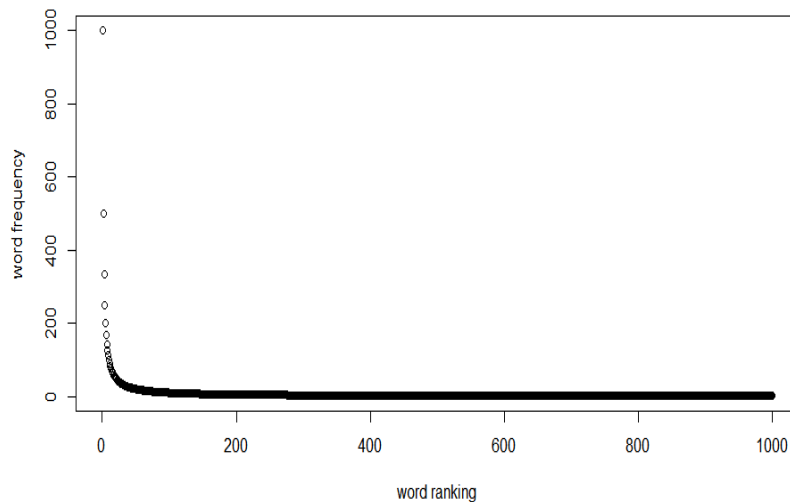
939 |EUNMI LEE

how to apply an innovative method to analyze a policy study exploring zipf's law on
Korean economic policy direction reports

Figure 1. *Example of Zipf Distribution (1st rank word frequency based on 1000)*

In the power distribution, large-scale phenomena rarely appear, and small-scale phenomena appear as non-linear shapes that frequently occur. This distribution is observed in various social phenomena such as word frequency, paper citation frequency, earthquake size, city population, war intensity, business size, and housing price [11,57]. In particular, in the policy document, a Zipf distribution with an inverse regularity between word ranking and word frequency appears. Figure 1 is an example of the Zipf distribution, assuming that the first-order word frequency is 1,000 and shows the graph in the form of "L."The Zipf distribution, like the power function, has only a few words with a high frequency, and most words have a long tail structure with a low frequency. In the Jeep distribution, the scale, such as the mean or median, cannot explain the whole, but it has a self-similarity that repeatedly replicates the pattern of an "L" shape regardless of the document size.

The fourth argument is related to constraints in the Zipf distribution, and the policy document consists of a dual structure of a small number of high-frequency words that are continuously selected at a high frequency and many rare-frequency words that are frequently occurring under a historical context and circumstance. Moreover, the middle area where the high-frequency words and the rare-frequency words overlap is relatively narrow as shown in Figure 1.A small number of high-frequency words in the form of "I" in Figure 1 are duplicated and replaced each time to form the basis of the policy document and maintain the continuity and stability of the policy document. High-frequency words do not change easily because they reflect the social perception of understanding the phenomenon and represent the norms of the governing order, and accordingly, the policy document also has stable attributes. However, many rare-frequency words corresponding to the long tail of "—" in Figure 1 are replaced with different words each time to create variability and differentiation in policy documents. The short-term and adaptive changes of policy documents are determined according to the differences in the rare-frequency words of each document. This characteristic of the Zipf distribution creates the duality of policy changes and creates a kind of optical illusion. For example, when the policy document changes by 30%, this does not mean that the overall 30% is changed uniformly every time but that the replacement occurs in 30% of the area where the constraint is weak and that the remaining 70% of the area remains unchanged. It looks like 30% of the time, it has changed each time, but the rest of the change is maintained, whereas the rest changes every time. In the short term, variability in policy documents is produced to reflect the preferences and willingness of policy actors, but in the long term, high-frequency words under the constraints of the institutions are buffering rapid policy changes.

Fifth, as a policy document on the mechanisms of long-term change is transformed into "inversions" in "alternative" overlapping areas, it replaces the scarce frequency words that correspond to the long tail as a new word. Rare-frequency words appear opportunistically according to the needs of the time, and then, are replaced voluntarily with other words. The more that the policy actors use new and rare words that are not adopted by other documents, the more that the policy documents are changed and that the more local and temporary changes are created. However, rare words are replaced by other words without being replaced by the next policy document. However, policy documents will change in the long term through a change in overlapping areas. Some of the newly transferred words move to the overlapping area in accordance with social problem recognition and the national mood [19], and some of the existing high-frequency words are shrunk and intersect with the words in the overlapping area. Changes in the overlapping domain will be influenced by the social acceptance of problem-solving

solutions, new policy learning to achieve goals, the transformation of policy alternatives according to social events, governmental exchanges, and the transition of state affairs. As shown in Figure 1, due to the nature of the Zipf distribution, the overlapping area is not structurally wide, and accordingly, the change through inversion is inevitable.

In short, the dual structure of the Zipf distribution, reflecting the historical context and the institutional situation, creates policy stability and variability simultaneously. High-frequency words subject to strong institutional constraints are used repeatedly and in conjunction to form the basic trajectory and framework of policy documents and lead to path-dependent stability. Rare-frequency words [59,60,61,62,63], which are discretionary choices by policy actors under loose constraints, create policy variability, but such changes will be limited and temporary. Furthermore, a long-term change occurs due to word substitution in the overlapping area where rare and high-frequency words intersect, but the scope of the change is narrow and trivial. The next section explains the details of how these discussions are reflected in actual cases by analyzing the economic policy direction report, which is a document produced by the government.

EMPIRICAL RESULTS

To demonstrate policy changes under institutional constraints through the Zipf distribution of policy documents, the 25-year economic policy direction report from 1993 to 2017 was decomposed and analyzed in words. The economic policy direction announced by the government at the beginning of each year is a combination of policies and economic policy goals to be pursued during the year and which present major policy tasks. The economic policy direction as an analysis object has been regularly announced since the 1960s, so it has the advantage of easy comparative analysis. This is different from the ministry plan, which makes it difficult to guarantee long-term work consistency due to government reorganization.

To transform the economic policy direction into algebraic numerical data in vector space, this study extracts words from the economic policy direction using the R package. Because the economic policy direction has a noun-centered revised form in which investigation and endings are omitted, the words are separated based on spaces. Moreover, the compound word composed of several words was processed by referring to the Current Economics Dictionary of the Ministry of Strategy and Finance and the Economic Terminology Dictionary of the Bank of Korea. Furthermore, numbers, punctuation, stop words, and one-syllable words were removed, and English and Korean with the same meaning were organized into Korean words. Also, if abbreviations are mixed, such as in local governments, joint commissions, and joint committees, they are corrected with a uniform word and processed. After this pre-processing, a vector space was created around the index word, which acts as an identifier. The total number of syllables was 349,430, the number of words was 81,679, and the index word was 9,364.
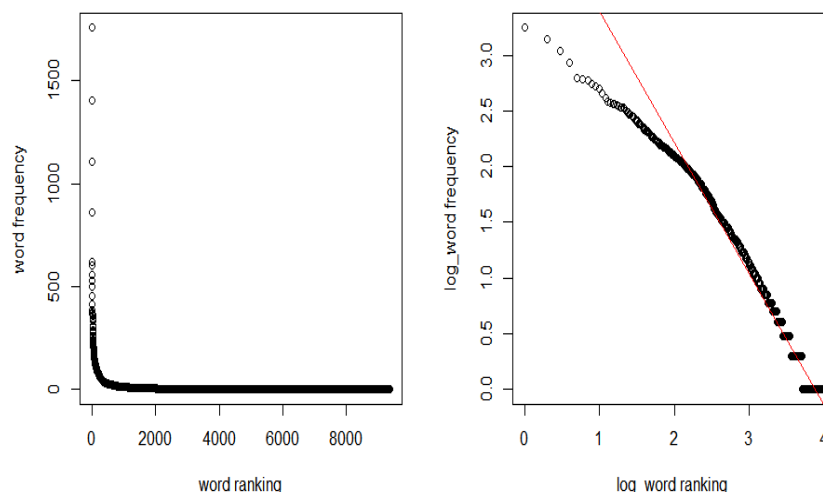

Figure 2. *Word ranking and frequency graph/log word ranking and log frequency graph*

Accordingly, to examine whether the economic policy direction corresponds to the dual structure of the Zipf distribution while simultaneously having variability and stability, the word rank and word frequency are described as the X and Y axes, respectively. As shown in the graph on the left side of Figure 2, the word distribution in the direction of economic policy is expressed as an inversely non-linear decreasing curve with a long tail of the "L" type. The word used means that it appears very often.

The most frequent words in the graph are "support" (1,755 times), followed by "expansion" (1,493), "enhancement" (1,104), "improvement" (863), and "enterprise" (619). Based on the frequency, 1,873 words corresponding to 20% of the index words were used 67,860 times, constituting 83.1% of all words. Reconstructing the direction of economic policy based on frequently appeared words, the direction of economic policy is to support companies, markets, finance, and industries to expand, strengthen, promote, enhance, and revitalize existing policies and to introduce, develop, and introduce new measures. Specifically, the economic policy direction based on "supporting" economic entities, such as markets and companies, is based on social awareness of "markets and enterprises, growth and welfare, traditional and new industries, labor markets, inflation and prices, international trade," etc. Several words that reflect economic norms are presented annually in the form of rearrangement and relocation. This aims to maintain the continuity of the past and present while the existing words constituting the economic policy direction remain under institutional constraints and newly combine with new words.

Meanwhile, the number of words used fewer than five times was 7,345 or 78.4% of the index words and 15.8% of all words. These words are rare-frequency words corresponding to long tails and serve to differentiate the economic policy direction of each time. Policy actors are able to respond appropriately to the policy environment and cope with external pressures for new things by describing the scarce frequency words of the discretionary policy options under loose institutional constraints.

To statistically reaffirm whether the long-tailed "L"-type graph is a Zipf distribution, a log was taken for the word rank and word frequency values. As shown in the graph on the right in Figure 2, it will be converted into a straight line that decreases in the downward direction. In the case of a perfect Zipf distribution that is inversely proportional to the word rank, the slope value of the regression equation is 1, but if the empirical coefficient value is between 0.8 and 1.2, it is considered to belong to the Zipf distribution [64]. According to these criteria, the slope value of the word frequency distribution is 1.18, and the economic policy direction to be analyzed is said to correspond to the Zipf distribution with the dual structure of high and rare-frequency words.

Table 1. *Statistical Values*

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	4.580186	0.007787	588.1	<2e-16 ***
log_rank	-1.180378	0.002185	-540.2	<2e-16 ***
Signif. codes:	0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			
Multiple R-squared:	0.9689, Adjusted R-squared: 0.9689			
F-statistic:	2.918e+05 on 1 and 9362 DF, p-value: < 2.2e-16			

To examine the dual structure of the Zipf distribution from the other side, this study used the document frequency, which is the number of times the word appeared. As shown in the graph in Figure 3 below, 36 index words were used in all 25 economic policy directions reports, and these words were used 14,865 times, corresponding to 18.1% of all words. Moreover, the index word used in 21–25 documents was 148, which was used 30,679 times, making up 37.6% of the entire word. By repeating and replacing each of these words, it plays a role in maintaining the basic framework of economic policy direction and continuity. However, due to the frequently appearing characteristics, the value of the information embedded in the word is not relatively high, and it cannot be said to contribute to the document's differentiation.
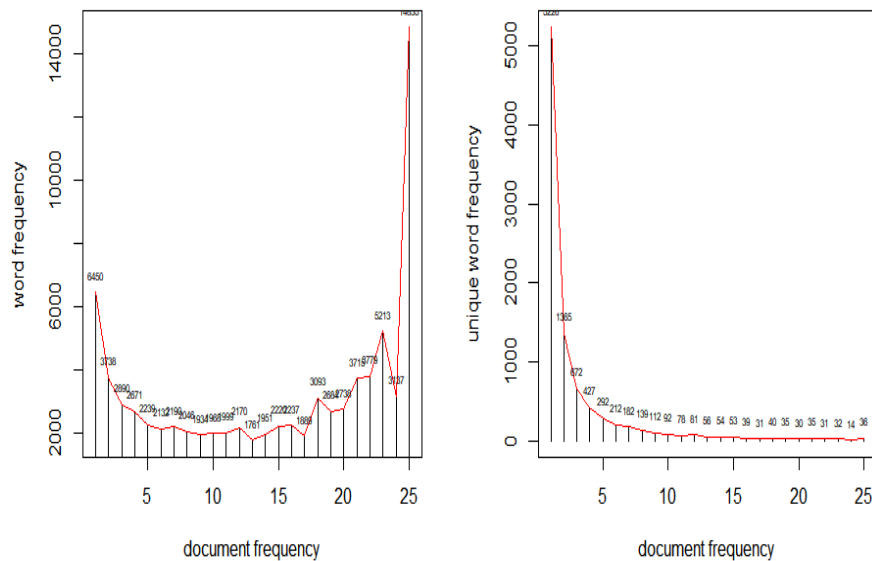
Figure 3. *Document frequency and word frequency graph/Document frequency and index word frequency graph*

However, as shown in the graph on the right in Figure 3, 5,226 index words appeared in only one document and 7,982 index words appeared in 1–5 documents. The rare-frequency words appearing at such a low frequency accounted for 85.2% of the index words but only 22.0% of the total words. In Figure 3, the words appearing in 6–20 documents, which can be called the middle zone, showed a relatively stable appearance. The index words used in 6–20 documents was 1,234, which was 13.2% of the index words. The change in the middle area has a mode of being maintained gently and stably, which reflects that the overlapping portion of the high and rare-frequency words is not large under the dual structure of the Zipf distribution.

CONCLUSIONS

As described above, this study examined how the dual structure of policy documents creates variability and stability of policy changes in connection with institutional research for the economic policy direction produced by the government.

In the analytical results, the economic policy direction showed a variability that changes more than a certain level each time, but it represented a modest change in the long run. The dual structure of the word frequency of the Zipf distribution is intended to reflect both the variability and stability of policy changes. The rare-frequency words used in the economic policy direction produced temporary and local changes in the short term, allowing the economic policy direction to adapt flexibly to the policy environment and political situation. Rare-frequency words, under loose constraints, are easily adopted by policy actors without causing major backlash because they do not directly impede social perceptions and dominant interests in the economic policy direction.

Yet, high-frequency words, which are strongly subject to institutional constraints, kept the economic policy direction stable while suppressing rapid changes in policies. High-frequency words that appear frequently every time serve as a copy of the previous economic policy direction and an original volume of future copies. The word change in the overlapping area, where the rare and high-frequency words intersect is made smoothly under the influence of the change of national keynote, governmental change, social acceptance of new problem-solving solutions, and transformation of policy alternatives according to social events. This served to create policy changes over time. In summary, as a characteristic of the dual structure of the Zipf distribution, the variability and stability of policy changes were simultaneously expressed in policy documents. In particular, the rare-frequency words represent the variability to cope with the policy environment, and the high-frequency words are the social norm, value, and shared meaning o institutions to reflect the stability of the policy.

This study examines a kind of contradictory phenomenon of variability and stability of policy changes by using vector space models and Zipf distribution for policy documents. To converge policy change and institutional research, it is intended to interpret policy changes with a new approach that is different from existing research methods. However, the limitations and discussions that are exposed in the process need to be reinforced in future studies. First, there should be a shared consensus among

943 |EUNMI LEE

how to apply an innovative method to analyze a policy study exploring zipf's law on
Korean economic policy direction reports

researchers as to whether policy changes can be understood through word changes in policy documents. This is a question underlying text analysis, which will require continued support from the academic community and additional methodological support.

Second, this study analyzed the degree of change through word frequency based on the Zipf distribution, but supplementary and additional case analysis is required to validate this reasoning. Researchers should also discuss whether the Zipf distribution presented in terms of experience can be directly applied to generalized discussions about policy changes. Third, the vector space calculates only the frequency without considering the order in which the words appear, but in the future methodology, it is necessary to analyze in what order the words are combined and arranged. Researchers should also develop techniques for processing natural language from a word-level analysis to a sentence-level analysis. Fourth, this study focuses on the unilateral effect of institutional constraints on policy changes, and it does not highlight that the policy not only is influenced by the institution but also has an active character that affects the institution. Through this methodological supplementation and discussion in the future, this study expects that an alternative explanation for policy change can be reinforced and that the policy theory can be expanded to many fields.

Funding

Conflicts of Interest

ACKNOWLEDGEMENTS

REFERENCES

March, James G., and Johan P. Olsen. (1984). The new institutionalism: Organizational factors in political life. *American Political Science Review, 78*(3), 734-749.

March, James G., and Johan P. Olsen. (1989). *Rediscovering Institutions: The Organizational Basis of Politics.* New York: The Free Press.

Hall, Peter A. (1986). *Governing the Economy: The Politics of State Intervention in Britain and France.* Oxford University Press.

Thelen, Kathleen, & Sven Steinmo. (1992). Historical institutionalism in comparative politics. In Sven Steinmo, Kathleen Thelen and Frank Longstreth(eds.), Structuring politics: historical institutionalism in comparative analysis, 369-404. Cambridge University Press.

Ikenberry, G. John. (1988). Conclusion: an institutional approach to American foreign economic policy. *International organization, 42*(1), 219-243.

Hall, Peter A., and Rosemary CR Taylor. (1996). Political science and the three new institutionalisms. *Political studies, 44*(5), 936-957.

North, Douglass C. (2005). *Understanding the process of institutional change.* Princeton, NJ: Princeton University Press.

Steinmo, Sven, Kathleen Ann Thelen, and Frank Longstreth. (1992). *Structuring politics: historical institutionalism in comparative analysis.* Cambridge University Press.

Immergut, Ellen M. (2006). Institutional Constraints on Policy. In Michael Moran, Martin Rein and Robert E. Goodin(eds.). *The Oxford Handbook of Public Policy,* 557-571. Oxford: Oxford University Press.

Thelen, Kathleen. (1999). Historical institutionalism in comparative politics. *Annual review of political science, 2*(1), 369-404.

Yun, J.J., Jeong, E., & Yang, J. (2015). Open innovation of knowledge cities. *Journal of Open Innovation: Technology, Market, and Complexity*, *1*(2), 16.

Pontusson, Jonas. (1995). From comparative public policy to political economy: Putting political institutions in their place and taking interests seriously. *Comparative political studies, 28*(1), 117-147.

Sabatier, Paul A., and Hank C. Jenkins-Smith. (1993). *Policy change and learning: An advocacy coalition approach.* Boulder, Colo.: Westview.

Birkland, Thomas A. (1998). Focusing events, mobilization, and agenda setting. *Journal of Public Policy, 18*(3), 53-74.

Birkland, Thomas A. (2006). *Lessons of Disaster: Policy Change after Catastrophic Events.* Georgetown University Press.

Hall, Peter A. (1993). Policy paradigms, social learning, and the state: the case of economic policymaking in Britain. *Comparative politics, 25*(3), 275-296.

May, Peter J. (1992). Policy learning and failure. *Journal of public policy, 12*(4), 331-354.

Wincott, Daniel. (2011). Ideas, policy change, and the welfare state. In Daniel Béland and Robert Henry Cox(eds.), *Ideas and politics in social science research,* 143-166. Oxford University Press.

Kingdon, John W. (1995). *Agendas, Alternatives and Public Policies,* 2nd Ed. New York: Haper Collins.

Atkinson, Michael M. (2011). Lindblom's lament: Incrementalism and the persistent pull of the status quo. *Policy and Society, 30*(1), 9-18.

Lindblom, Charles E. (1959). The Science of 'Muddling Through'. *Public Administration Review, 19*(2): 79-88.

Lindblom, Charles E. (1964). Contexts for Change and Strategy: A Reply. *Public Administration Review, 24*(3): 157-58.

Immergut, Ellen M., and Karen M. Anderson. (2008). Historical institutionalism and west European politics. *West European Politics, 31*(1-2), 345-369.

Ragin, Charles C. (1987). *The comparative method: Moving beyond qualitative and quantitative methods.* Berkeley: University of California.

Skocpol, Theda. (2003). Doubly engaged social science: The promise of comparative historical analysis. in Mahoney, James, and Dietrich Rueschemeyer (eds). *Comparative Historical Analysis in the Social Sciences.* 407-28. Cambridge University Press.

Skocpol, Theda, and Margaret Somers. (1980). The uses of comparative history in macro social inquiry. *Comparative Studies in Society and History, 22*(2), 174-197.

Lindblom, Charles E. (1979). Still Muddling, Not Yet Through. *Public Administration Review, 39*(6): 517-26.

Simon, Herbert. (1945). *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations.* New York: The Free Press.

Lindblom, Charles E. (1977). *Politics and Markets: The World's Political-Economic Systems.* New York: Basic Books.

March, James G., and Johan P. Olsen. (1996). Institutional perspectives on political institutions. *Governance, 9*(3): 247-264.

March, James G., and Johan P. Olsen. (2004). *The logic of appropriateness.* The Oxford handbook of political science.

Hall, Peter A. (1992). The movement from Keynesianism to monetarism: Institutional analysis and British economic policy in the 1970s. In Sven Steinmo, Kathleen Thelen, and Frank Longstreth (eds.), *Structuring politics: Historical institutionalism in comparative analysis.* 90-113. New York: Cambridge Universtiy Press.

Gould, Stephen Jay, and Niles Eldredge. (1972). Punctuated equilibria: an alternative to phyletic gradualism. *In Essential readings in evolutionary biology,* 82-115.

Gould, S.J., & Eldredge, N. (1977). Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology, 3(2), 115-151.

Krasner, Stephen D. (1984). Approaches to the state: Alternative conceptions and historical dynamics. *Comparative Politics, 16*(2), 223-246.

Mahoney, James. (2000). Path dependence in historical sociology. *Theory and Society, 29*(4), 507-548.

Arthur, W. Brian. (1988). Self-reinforcing mechanisms in economics. *The Economy as an Evolving Complex System,* 9-31.

Arthur, W.B. (1994). *Increasing returns and path dependence in the economy.* University of michigan Press.

David, Paul A. (1985). Clio and the Economics of QWERTY. *The American Economic Review, 75*(2), 332-337.

North, Douglass C. (1990). *Institutions, Institutional Change and Economic Performance.* Cambridge University Press.

Pierson, Paul. (2000). Increasing returns, path dependence, and the study of politics. *American Political Science Review, 94*(2), 251-267.

Streeck, Wolfgang, and Kathleen Thelen. (2005). *Beyond continuity: Institutional change in advanced political economies.* Oxford University Press.

Thelen, Kathleen. (2004). *How institutions evolve: The political economy of skills in Germany, Britain, the United States, and Japan.* Cambridge University Press.

Zipf, George Kingsley. (1949). *Human behavior and the principle of least effort.* Hafner Publishing Company.

Blyth, Mark. (1997). Any more bright ideas? The ideational turn of comparative political economy. *Comparative Politics, 29*(2), 229-250.

Blyth, M., & Mark, B. (2002). *Great transformations: Economic ideas and institutional change in the twentieth century.* Cambridge University Press.

Ideas, uncertainty, and evolution. In Daniel Béland and Robert Henry Cox(eds.), *Ideas and politics in social science research,* 83-101. Oxford University Press.

Campbell, John L. (1998). Institutional analysis and the role of ideas in political economy. *Theory and Society, 27*(3), 377-409.

Mehta, Jal. (2011). The varied roles in politics: from whether to how. In Daniel Béland and Robert Henry Cox(eds.), *Ideas and politics in social science research,* 83-101. Oxford University Press.

Jordan, Meagan M. (2002). Unctuated Equilibrium: An Agenda-Based Theory of Budgeting. In A. Khan & W. B. Hildreth(eds.), *Budget Theory in the Public Sector,* 202-215. London: Quorum Books.

Wildavsky, Aaron B. (1964). *The Politics of the Budgetary Process.* Boston, MA: Little, Brown and Company.

Bailey, John J., and Robert J. O'Connor. (1975). Operationalizing incrementalism: Measuring the muddles. *Public Administration Review, 35*(1), 60-66.

Kemp, Kathleen, A. (1982). Instability in budgeting for federal regulatory agencies. *Social Science Quarterly, 63*(4), 345-360.

Baumgartner, Frank R., and Jones, Bryan D. (1993). *Agendas and Instability in American Politics.* Chicago: University of Chicago Press.

Jones, Bryan D., Baumgartner, Frank R., and True, James L. (1996). The Shape of Change: Punctuations and Stability in U.S. Budgeting, 1947-1994. Presented at the *Annual meeting of the Midwest Political Science Association.* Chicago.

Salton, Gerard, Anita Wong, and Chung-Shu Yang. (1975). A vector space model for automatic indexing. *Communications of the ACM, 18*(11), 613-620.

Gabaix, Xavier. (1999). Zipf's Law and the Growth of Cities. *American Economic Review, 89*(2), 129-132.

Mandelbrot, Benoit B. (1983). *The fractal geometry of nature.* Vol. 173. New York: WH freeman.

Ho, J.C., & Zhang, X. (2020). Strategies for Marketing Really New Products to the Mass Market: A Text Mining-Based Case Study of Virtual Reality Games. *Journal of Open Innovation: Technology, Market, and Complexity*, *6*(1), 1.

Saura, J.R., Debasa, F., & Reyes-Menendez, A. (2019). Does User Generated Content Characterize Millennials' Generation Behavior? Discussing the Relation between SNS and Open Innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(4), 96.

Saura, J.R., Reyes-Menendez, A., & Filipe, F. (2019). Comparing Data-Driven Methods for Extracting Knowledge from User Generated Content. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(4), 74.

Saura, J.R., Reyes-Menendez, A., & Palos-Sanchez, P. (2019). Are black friday deals worth it? Mining Twitter users' sentiment and behavior response. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(3), 58.

Lee, B., Park, J.H., Kwon, L., Moon, Y.H., Shin, Y., Kim, G., & Kim, H.J. (2018). About relationship between business text patterns and financial performance in corporate data. *Journal of Open Innovation: Technology, Market, and Complexity*, *4*(1), 3.

Nitsch, Volker. (2005). Zipf zipped. *Journal of Urban Economics, 57*(1), 86-100.

Dye, Thomas R. (2001). *Top down policymaking.* NY: Chatham House Pub.

Huang, Anna. (2008). Similarity measures for text document clustering. *Proceedings of the sixth New Zealand computer science research student conference (NZCSRSC2008),* Christchurch, New Zealand. Vol. 4.

Salton, Gerard, and Christopher Buckley. (1988). Term-weighting approaches in automatic text retrieval. *Information processing & management, 24*(5), 513-523.