A Study of Text mining for Research into the History of Economic Thought:

The case of Alfred Marshall's *Principles of Economics* (1890)

Naoki Matsuyama

School of Economics
University of Hyogo

February   2016

A Study of Text mining for Research into the History of Economic Thought:

The case of Alfred Marshall's *Principles of Economics* (1890)*

Naoki Matsuyama, M.A., Ph.D.

School of Economics, University of Hyogo, Japan

naoki.ma@econ.u-hyogo.ac.jp

Abstract

This paper investigates whether text mining can develop the general features of a particular work of economic literature, Alfred Marshall's *Principles of Economics* (1890), with a minimum of subjectivity and arbitrariness. Text mining has been receiving more attention from historians of economics during recent years. In this paper, text-mining analysis shows that the general features of Marshall's economics are closely related to Book V and VII of his work. Consequently, this research clarifies that text mining could be employed as a supplementary research method in studying the history of economic thought.

Keywords: Text mining, the history of economic thought, Alfred Marshall, *Principles of Economics*, research method

1. Introduction

This paper provides a brief case study of text mining applied to hermeneutic study into the history of economic thought. In particular, I investigate whether text mining can develop the general features of a particular work of economic literature, Alfred Marshall's *Principles of Economics* (1st edition, 1890, hereafter *Principles*). Text mining is a branch of data mining in computer science and as a quantitative research method well-equipped for philological research[1]. This study therefore provides the idea on text mining to extract objective and unbiased information from the first edition of *Principles*[2].

---

[1] Tokosumi and Murai (2014) demonstrate many examples of text mining applied to philological studies, such as literature, history, and music.
[2] This research has a root in studying ontology. In fact, the study of ontology relates not only to philosophy but also to both economics (Lawson 1997, Martins 2014) and computer science (Smith 2003). This research trend in ontological studies has greatly affected the basic idea of this study, especially the limit and potential of text mining applied to the history of economic thought.

In recent years, text mining has been receiving more attention from historians of economics. They have attempted to show the dissemination process of an economic doctrine through some economic literature (Shimodaira *et al.* 2012; Furuya 2014; Shimodaira and Fukuda 2014). A few econophysicists have also tried to capture the essential features of economic literature using a statistical approach (Trincado and Vindel 2015). These earlier studies have been intended not only to add objective evidences to a historian's subjective interpretation, but also to rediscover any buried ideas historians of economics might have missed in the huge number of textbooks and manuscripts.

As far as I understand it, however, previous studies are still insufficient for use in text mining as scientific analysis suitable for the study in the history of economic thought. That is to say, is the research impact of text mining in this field restricted to only express the dissemination process of an important notion through some economic literature? It should be necessary at this point to investigate whether text mining can capture any quantitative features of *a single work* in economic literature that harmonizes with hermeneutic research into the history of economic thought. This study offers subjective and objective aspects of the first edition of *Principles* using text mining. As a result, text mining can be seen as a viable research method in the historical study in economics.

This paper, therefore, investigates the possibilities of applying text mining to the study of the history of economic thought. Further, this study could play a key role in demonstrating whether text mining can extract general features of Marshall's economics, from the first edition of *Principles*, with a minimum of subjectivity and arbitrariness.

2. Subject of Research and Methods

2-1. Subject of Research

To confirm the application of text mining to the history of economic thought, the subject of this paper is the first edition of Alfred Marshall's *Principles*, published by Macmillan and Co. Ltd in London in 1890. *Principles* has attracted positive attention from readers over a long period. It had been used as a standard economics textbook all over the world until World War II. This paper attempts to capture the main arguments of *Principles* by means of text mining.

Traditionally, *Principles* has been considered a basic textbook of economic analysis. In the preface, Marshall indicates simply the main role of economic analysis as follows:

> Economic laws and reasonings* in fact are merely a part of the material which Conscience and Common-sense have to turn to account in solving practical problems, and in laying down rules which may be a guide in life.
>
> Marshall 1890, p.vi *without any change

Marshall believes that economic analysis makes only a partial contribution to the problem-solving process of our daily lives. Therefore, '[p]olitical economy, or economics, is a study of man's actions in the ordinary business of life', Marshall maintained in the first paragraph of his *Principles*. '[I]t is on the one side a study of wealth and on the other, a more important side, a part of the study of man' (Marshall 1890, 1).

Marshall's understanding of economic analysis has been widely known to contemporaries, through not only his reputation as an economics professor but also through several book reviews in major academic journals. For instance, Francis Ysidro Edgeworth (1845−1926) published a review of Marshall's *Principles* in *Nature* when he was Professor of Political Economy at King's College, London. Edgeworth correctly indicated the methodological foundations of Marshall's *Principles* as follows:

> He [Prof. Marshall], of all mathematical economists, has best compiled with his own maxim that the economist, while he employs 'systematic reasoning as to the quantities of measureable motives, ... must never lose sight of the real issue of life and these are all, with scarcely any important exceptions, affected more or less by motives that are not measurable'.
>
> Edgeworth [1890]1998, 8

Thus, as Edgeworth clarified, Marshall's *Principles* contains two methodological approaches. The first is the mechanical reasoning, as expressed by mathematics and graphical representation, in order to statically grasp economic phenomena. The second is the practical and moral thinking around economic problems, such as poverty. Since the latter cannot be easily understood by

mathematical expression, Marshall always emphasizes that economists should closely observe real life and attempt practically to elucidate complicated economic phenomenon. However, he could not develop the framework of the latter to the same level as the equilibrium theory of demand and supply in his lifetime.

How does text mining approach the first edition of *Principles* quantitatively? Fortunately, since *Principles* is a book in which, not like French economics books, mathematical expressions are restricted as much as possible in the main text, making it more suitable for text mining. The structure of the first edition is as follows:

Table of Contents, Marshall's *Principles of Economics* (1890)

  Book I. — Preliminary Survery. (pp. 1−98)

  Book II. — Some Fundamental Notions. (pp. 101−144)

  Book III. — Demand or Consumption. (pp. 147−183)

  Book IV. — Production or Supply. (pp. 187−380)

  Book V. — The Theory of the Equilibrium of Demand and Supply. (pp. 383−478)

  Book VI. — Cost of Production Further Considered. (pp. 481−536)

  Book VII. — Value, or Distribution and Exchange. (pp. 539−736)

  Mathematical Appendix (pp. 737−750)

  Index (pp. 751−754)

2.2. Method

This section explains the detailed research method of this paper. Text mining is a method of automated processing text and 'an approach to doing research that begins with words rather than numbers', as Miller (2005, 104) observed. Further, text mining is incomplete without very careful preprocessing. Unless anyone does not follow the precise preprocessing, the software of text mining cannot correctly compile and count every instance of words without false positive. The basic procedure, including the preprocessing this paper has adopted, is composed of the following five steps:

(1) To prepare a PDF file of the original text of *Principles*[3]

---

[3] In 2014, the British government passed Regulation 2014, No.1372, to the effect that text mining for non-commercial research is not in conflict with copyright law. See also *The Copyright and Rights in*

(2) To transform the PDF file for each unit of analysis into Microsoft Word format

(3) To correct unreadable characters, and to modify errata, such as Latin expressions

(4) To convert the revised Microsoft Word format text format

(5) To analyse the text file of *Principles* with the text-mining software KH Coder[4]

In this regard, this study has applied text mining only to the body of *Principles*, from Book I to Book VII, and does not include footnotes, notes, or Appendices. Furthermore, this research has focused on *nouns* only − a cognitive psychology study has mentioned that 'noun shows an idea, a notion and a philosophy most symbolically' (Kida 2008, 151).

This study employs three representative methods of text mining by means of KH Coder: frequency analysis, co-occurrence network analysis, and correspondence analysis. Frequency analysis is the most basic approach to investigating the frequency of words used. In fact, this method has been very familiar within the study of the history of economic thought since the mid-twentieth century (Matsukawa 1969; Gherity 1993; Sudo 1999). This analysis shows that a high/low frequency denotes high/low importance of the term. Co-occurrence network analysis is one of the best known visualization methods to develop networks of related terms. Each network shows the degree of strength for the co-occurrence of related terms, and then a highly used-frequency term is drawn with a big circle in the network figure. This method is divided into two specific components: centrality analysis and community analysis. The former is a method of identifying how central a role each term plays in the network figure. The colour of each circle − sky blue, white and pink - designate degrees of centrality. The latter is a method to show communities with different colours with which each term comparatively has a strong relationship. In this regard, all terms included in the same community are connected by solid lines. Finally, correspondence analysis is a method of visualizing the distribution of terms as structural data in the textbook. In this structural data, all key terms can be characterized by similar frequency patterns and gathered together with each of the components of the textbook. This paper employs the Bubble plot function of KH Coder in order to show the distribution

---

Performance (Research Education, Libraries and Archives).
URL: http://www.legislation.gov.uk/uksi/2014/1372/regulation/3/made
[4] KH Coder is freeware. The central purpose of this paper is to investigate whether or not the text-mining method can extract some theoretical features from a copy of *Principles*. Therefore, this study does not discuss the statistical structure of KH Coder.
Regarding this issue, see Higuchi (2014) URL: http://sourceforge.net/projects/khc/

position for each component, from Book I to Book VII, in the structural data of the first edition of *Principles*.

3. Result and Discussion

The following two levels of the text-mining method that this paper adopts enable us to acquire the general features of the first edition of *Principles*. In the first level, this study employs frequency analysis and co-occurrence network analysis in order to determine significant terms of the first edition of *Principles*. In the second level, I demonstrate the distributions of all key terms brought out by the first step into the structural data of *Principles* by means of correspondence analysis.

3-1. On the General Features of *Principles*

Text-mining analysis could quantitatively identify the important contents of *Principles* as the general features. First, the results of co-occurrence network analysis such as Figures 2 and 3, demonstrate that *Principles* is constituted of three major communities, led by the following three nouns: 'capital', 'labour', and 'production'. Further, it is clear that these three nouns play a key role in *Principles* and have their own communities, according to frequency analysis and centrality analysis: the community of 'capital' (hereafter the Capital Community), the community of 'labour' (hereafter the Labour Community), and the community of 'production' (hereafter the Production Community). Also, co-occurrence network analysis suggests that the most frequently used noun is 'price', and 'price' is one of the components of the Production Community. Therefore, the Production Community could have a higher degree of importance than either the Capital Community or the Labour Community, from quantitative analysis of *Principles*. The detailed components of each community as the general features of *Principles* can be confirmed in Table 1.

Second, as Figure 1 shows, correspondence analysis indicates the main theme of each Book of the first edition of *Principles*, and this identifies each attribution of the three communities[5] in *Principles*. In particular, the clearest result of correspondence analysis was the independent

---

[5] As I mentioned in Table 1, (1) the Capital Community corresponded to Book II, Book IV, and Book VII, (2) the Labour Community was reflected by Book VII, and (3) the Production Community was reflected by Book V and Book VI of *Principles*.

Table 1. Brief summary of the results of text-mining applied to *Principles*

Community	Components of Community	Contents of "Principles"
Capital Community	(capital, business, man, ability)	Book II, Book IV, Book VII
Labour Community	(labour, kind, wages)	Book VII
Production Community	(production, demand, commodity, price…)	Book V, Book VI

* Each component of communities was placed in order of degree of frequency.

** The detailed results of text mining applied to *Principles* are in Table 3 of the Appendix.

Figure 1. The Results of Corresponding Analysis applied to *Principles*
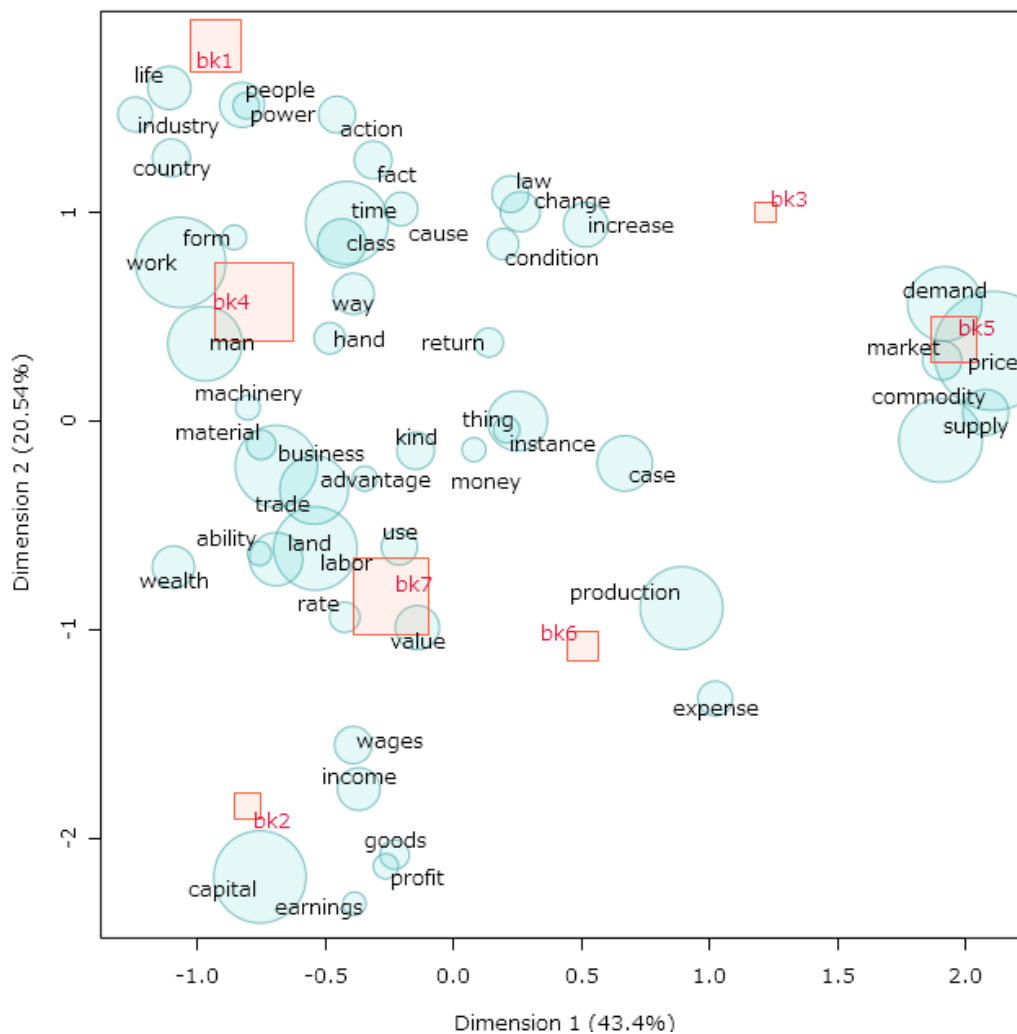
* Experimenters can correct many functions of KH Coder, including built-in dictionary. However, the author does not have any skills to arrange the setting of the built-in dictionary from American English to British English. Thus, the results of all figures in this paper are demonstrated in American English.

position of Book V of *Principles*. Book V was dominated by 'price', 'supply', 'demand', 'commodity', and 'market', which were closely related with the main components of the Production Community of *Principles*. In addition, Book IV and Book VII had the closest relationship in the first edition of *Principles*. The nature of this relationship was mediated by 'business' and 'trade', which were located midway in both Book IV and Book VII in Figure 1: however, in this analysis Book VII included the Labour Community, which was one of the general features of *Principles*, as I have already mentioned. Thus, the quantitative analysis suggests that Book VII describes general features of *Principles* more correctly than Book IV. Furthermore, it seems that correspondence analysis indicates that Book II has another important factor of *Principles* as the Capital Community. In this regard, however, there was no consistency between the results of co-occurrence network analysis and those of correspondence analysis. Therefore, the quantitative analysis suggests that Book II does not contribute significantly to describe the general features or of *Principles*.[6]

In short, text mining suggests that Book V and Book VII are significant in *Principles*, but it needs to read and interpret *Principles* in order to indicate and evaluate its essential point. In the next section, then, I identify the detailed and quantitative character of Book V and Book VII of *Principles*.

3-2. Two Specific Features of *Principles*: Book V and Book VII
Here, I begin more a specific and elaborate investigation into Book V and Book VII, with reference to Figures 8, 9, 12, and 13.

The main character of Book V lies in the Production Community (Figures 8 and 9). In particular, the term 'production' had the most central role in Book V, and the term 'price' was the most frequently used noun in Book V as well. These characters were mostly related to the Production Community, as one of the general features of *Principles* (Table 2)

Therefore, the quantitative character of Book V clearly emphasizes theoretical aspects such as supply and demand theory of *Principles*, through the properties of the Production Community.

---

[6] Conversely, it is possible that the Capital Community was one of the general features of *Principles*.

Table 2. The Production Communities

- On *Principles*	
Components:	'production' (506), 'price' (806), 'demand' (447), 'thing' (352), 'commodity' (274), 'increase' (266), 'expense' (217)
- On Book V	
Components:	'production' (150), 'price' (353), 'supply' (254), 'demand' (190),'commodity' (127), 'increase' (71), 'law' (52), 'schedule' (41), 'factor' (45)

* Each number shown in parentheses expressed the degree of frequency in use.

** This table is based on Figure 2 and Figure 9 in Appendix.

Second, according to the previous section, Book VII was the representative component of *Principles* in relation to the Labour Community[7]: the term 'labour' plays the central role in Book VII (Figure 12). Further, text mining suggests that the terms 'production' and 'capital' also form communities in Book VII[8] (Figure 13). As is the case with the general features of *Principles*, the Production Community in Book VII also includes the term 'price', which is the highest frequency noun in this community. Thus, all three nouns not only are in each community but are among the highest frequency of usage nouns in Book VII[9]. Therefore, in terms of quantitative analysis, it is clear that Book VII can be seen as an epitome of *Principles*.

In sum, Book V and Book VII appear to be central to the first edition of *Principles*: Book V focused on the ideas in the Production Community, which were the most important component of Marshall's *Principles*. Book VII resembles a miniature version of *Principles* as it encompasses the main communities found in *Principles* comprehensively. Hence, under the standpoint of text-mining analysis, Book V and Book VII should be recognized as the representative components of *Principles*. However, are these results acceptable for a correct understanding of Marshall's economics traditionally? It require to read and interpret them in the next step.

---

[7] The components of the Labour Community in Book VII are 'labour', 'kind' and 'grade'.

[8] The components of the Production Community in Book VII are 'production', 'price', 'supply', 'demand', 'commodity', 'agent', 'material' and 'thing'. The components of the Capital Community in Book VII are 'capital', 'business', 'profit', and 'management'.

[9] The top ten nouns in are 'capital' (249), 'labour' (240), 'business' (184), 'trade' (172), 'price' (169), 'wages' (167), 'work' (167), 'man' (147), 'earnings' (142), and 'production' (137), according to frequency analysis applied to Book VII.

3-3. Historical Understanding of Book V and Book VII

The main arguments of Book V has not only been placed as the core component of *Principles* but has also been recognized as the basis of modern economic theory. On the other hand, Book VII has been largely ignored by economists, who have paid much greater attention to the mathematical development of economic theory. Here, I confirm the general understanding of Book V and Book VII of *Principles*.

In Book V, Marshall has developed his well-known economic theory of the price mechanism, partial equilibrium theory, which was an analytical method of visualizing the interaction of the forces of demand and supply in a single market. In general, it is considered that classical economics before the emergence of Marshall's economics was based on Say's Law, in which the quantity of production automatically fulfilled the quantity of demand in a market. As a matter of fact, however, Marshall did not criticize this doctrine clearly, unlike John Maynard Keynes (1883−1946). Book V of *Principles* mainly discussed 'the price mechanism in a single market'. In particular, Marshall employed the term 'price' most frequently in his economic doctrine in Book V of *Principles*, according to text-mining analysis.

However, in fact, text mining failed to extract another of Marshall's contributions expressed in Book V, the concept of 'time' in his economic analysis[10]. Pigou (1907) clearly indicated that 'time' was one of the most important of Marshall's contributions to the development of economics[11]. Marshall himself emphasized the importance of the principles of continuity related with the concept of 'time' in the preface of his *Principles*. Furthermore, the concept of 'time' has been argued without major revision from the first edition to the last edition of *Principles* published in 1920.

---

[10] Marshall introduced four kinds of time concepts into his economic theory: temporary, short period, long period, and secular. The first three were used in static analysis. The last one was applied to analysis of economic progress, because the notion of secular permitted to change the number of population, technological level, and the scale of industry in a country. In addition, the latter need biological analogy based on evolutionary idea, not mechanical analogy based on mathematics. Therefore, Marshall maintained that 'the Mecca of the economists lies in economic biology rather than in economic dynamics' in the preface of *Principles*. Further, this division of time are related with his early research of psychology. See Matsuyama (2010).

[11] In fact, Book V has been combined with Book VI since the second edition of *Principles*. Pigou (1907) reviewed the fifth edition of *Principles*. His indication of the notion of 'time' is still effective and, in terms of the framework of Marshall's economic analysis, there has been no major revision in the contents of Book V from the first edition to the last.

Thus, in terms of the price mechanism, text mining could mostly capture the theoretical features of Book V. However, the fact related with the concept of 'time' tells us that text mining cannot always be the perfect analytical method.

Book VII, comprising thirteen chapters, covers the comprehensive ideas on Marshall's economics of the distribution and exchange in ordinary business of life. However, Book VII was merged into Book VI, with major revisions since the second edition of *Principles*, and Book VI had a major revision between the fourth edition (1898) and the fifth edition (1907). Most discussions of the fifth edition have focused on small revisions, until the last edition published in 1920. In terms of Book VI, especially its latter part, there were quite important arguments in terms of Marshall's ideas of economic progress. In particular, it is well known that Marshall's two original concepts, 'the standard of life' and 'economic chivalry', are recognized as the core ideas of his theory of economic progress. On the other hand, it is quite hard to trace such discussions in the first edition's Book VII. In fact, it seems that Book VII dealt with only the distribution and exchange related to capital and labour in order to develop the theory of economic progress in the later editions.

The main arguments of the latter part of Book VII were on the theory of national dividend, which was based the comparative history of American and British Industry. In fact, Marshall visited the United States of America in order to survey the conditions of protectionism, in 1875; he realized the importance of localized industries[12]. On the other hand, Marshall recognized that the UK economy in the nineteenth century was based not only on the development of transport industries, which have cut down 'the cost of transport of men and goods, of water and light, of electricity and news' (Marshall 1890, 718), but also on improvement in education, which brought about 'the growth of general enlightenment and of a sense of responsibility towards the young' (Marshall 1890, 724), in order to handle the increase of population and the law of diminishing returns. In addition, Marshall maintained that 'opportunities which the new methods of business offer for the safe investment of small capitals' (Marshall 1890, 729) also contributed to reduce poverty in the UK. In this point, therefore, the results of text mining, such as Figure 13, do not necessarily give us a correct understanding of Marshall's intentions.

---

[12] Matsuyama (2011) mainly discusses on Marshall's American research trip in 1875.

As we have discussed, the representative features of Book VII, constituted by the Labour Community, the Production Community and the Capital Community, do not clearly include any evidence related to the comparative analysis of the American and the UK economies. However, it can be said that text mining captures the main arguments of Book VII: the Labour Community can be regarded as a reflection of the discussion about income and quasi-rent in ordinary business life; both the Production Community and the Capital Community can be considered as reflections of the discussions of business profit from the view point of the demand and supply analysis. Also, the latter part of Book VII includes no evidence of his theory of economic progress, such as 'standard of life' and 'economic chivalry' expressed in Book VI of the later editions of *Principles*. Therefore, the results of text mining are relatively correct, but the quantitative method could not capture the peculiar and detailed part of Marshall's intentions embedded in *Principles*.

4. Conclusion

This study has tried to elucidate some quantitative features of the first edition of *Principles* by means of text mining: first, Book V includes Marshall's core theoretical argument of the demand and supply analysis; and second, Book VII develops comprehensive discussion of *Principles*. However, it is difficult for text mining to capture Marshall's specific and detailed intentions embedded in *Principles*. For instance, there are no discussions about the concept of 'time' in the results of the text mining applied to Book V. Indeed, some representative features extracted by text mining are based mainly on the frequency of usage of each noun. Text mining only expressed the degree of importance of each noun: the quantitative results of Book V show that 'time' is used just 57 times, while, in contrast, 'price' is used 353 times, 'supply' is used 254 times, and 'demand' is used 190 times. Thus, in order to capture the essential factors of *Principles*, intensive reading is absolutely required.

This study shows that the text-mining method is supplementary, but it is possible that text mining could be employed as a research method in studying the history of economic thought. According to previous studies, the significance of the text-mining method in this area lies in demonstrating the dissemination process of economic doctrines from one economic literature to others. Thus, the efficacy of the text mining-method can be confirmed in comparative analysis between economic literatures. In terms of the demonstrating text mining to capture some

features of a single economic literature, this paper plays an important role in the quantitative study of the history of economic thought.

This study brings us the following two perspectives. First, there is a space for the comparative study of the understanding of book reviews of the first edition of *Principles*, and its quantitative understanding is developed in this paper; it will be important for the study in the history of economic thought to replicate some contemporary researchers' understandings of the first edition of *Principles*. Second, this study brings us an opportunity to compare the quantitative features of different editions of *Principles*. As we discussed, it is easy to do comparative analysis between the first and the later editions of *Principles*. Therefore, future research will be focused on clarifying Marshall's inversion of his economic arguments, from the standpoints of both quantitative analysis and traditional qualitative analysis.

Finally, the characteristic arguments of Marshall's economics were unrestricted by abstract economic reasoning, such as the equilibrium theory discussed in Book V. Text-mining analysis more or less demonstrated the importance of the wider vision of practical economic activity discussed in Book VII. Therefore, the results of text mining in this research also teach us that economists and readers of economics literatures must be conscious of the relationship between practical economic activity and abstract economic doctrine.

References

Edgeworth, F.Y. [1890]1998. 'Book Review: *Principles of Economics*, Vol.I. By Prof Alfred Marshall. London: Macmillan and Co., 1890,' *Nature*, 14 August, 1890, in *Alfred Marshall: Critical Responses*, Vol.2, edited by P. Groenewegen. New York: Routledge, pp.3-8.

Furuya, Y. 2014. 'A Study on the Dissemination of Adam Smith's *Wealth of Nations* using Text-mining', *Discussion Paper, Tohoku Economics Research Group*, Tohoku University, No.325. [in Japanese]

Gherity, J.A. 1993. 'An Early Publication by Adam Smith', *History of Political Economy*, 25 (2), pp.241-282.

Higuchi, K. 2014. *Quantitative Text Analysis for Social Researchers*, Naganishiya-Shuppan. [in Japanese]

Kida, M. 2008. *An Introduction to Text-mining*, Hakusui-Shobo. [in Japanese]

Lawson, T. 1997. *Economics and Reality, Oxford*: Routledge.

Marshall, A. 1890. *Principles of Economics*, First Edition, London: Macmillan and Co.

Marshall, A. 1906. *Introduction to the Tripos in Economics and Associated Branches of Political Science*, Cambridge: Cambridge University Press.

Marshall, A. 1961. *Principles of Economics*, Variorum Edition, Vol.1 TEXT, London: Macmillan and Co Limited.

Matsukawa, S. 1969. '*Wealth of Nations* in N.E.D.', *Annals of the Society for the History of Economic Thought*, Vol.7, pp.85-87. [in Japanese]

Matsuyama, N. 2010. Relativity of Alfred Marshall's Psychological Research and Economics, *History of Economic Thought*, 51(2), pp. 52-67. [in Japanese]

Matsuyama, N. 2011. The Source of Marshall's Thoughts on Economic Progress with a Focus on his Study of American Industry, *Proceedings of the 24th Conference of the History of Economic Thought Society of Australia*, Royal Melbourne Institute of Technology, July 2011.

Martins, N.O. 2014. *The Cambridge Revival of Political Economy*, Oxford: Routledge.

Miller, T.W. 2005. *Data and Text Mining: A Business Applications Approach*, Pearson Prentice Hall.

Pigou, A.C. 1907. 'The Principles of Economics by Alfred Marshall', *Economic Journal*, Vol.17, No.78, pp.532-535.

Pigou, A.C. 1953. *Alfred Marshall and Current Thought*, London: Macmillan and Co LTD.

Shimodaira, H., Komine, A., and Matsuyama, N. 2012. 'An Introduction to Text-mining for the History of Economic Thought: Keynes' *General Theory* and its Book Reviews,' Discussion Paper, Research Group of Economics and Management, Yamagata University, No. 2012-E02. [in Japanese]

Shimodaira, H. and Fukuda, S. 2014. 'Popularization of Classical Economics: The text-mining Analysis of David Ricardo, James Mill, and Harriet Martineau', *Research Group of Economics and Management*, Yamagata University, No.2014-E01.

Smith, B. 2003. 'Ontology', in *Blackwell Guide to the Philosophy of Computing and Information*, Oxford: Blackwell, pp.155-166.

Sudo, Y. 1999. 'Adam Smith in OED', *The Hiyoshi review of English studies*, Keio University, Vol.35, pp.13-37. [in Japanese]

Tokosumi, A. and Murai, H. 2014. *Quantitative Analysis approach to Qualitative Research,*

Shinyo-Sha. [in Japanese]

Trincado, E. and Vindel, J.M. 2015. 'An application of econophysics to the history of economic thought: The analysis of texts from the frequency of appearance of key words', *Economics Discussion Papers* (Kiel Institute for the World Economy), No.2015-51.

Wood, J.C. ed. 1982. *Alfred Marshall: critical assessments*, Vol.I, II, III, IVm Kent: Croom Helm Ltd.

Wood, J.C. ed. 1996. *Alfred Marshall: critical assessments*, Second Series, Vol.V, VI, VII, VIII, New York: Routledge.

Figure 2. Network Figure of the first edition of Marshall's *Principles* (Centrality Analysis) (TN: 227,484, Min TF:160, NoSW:51)*
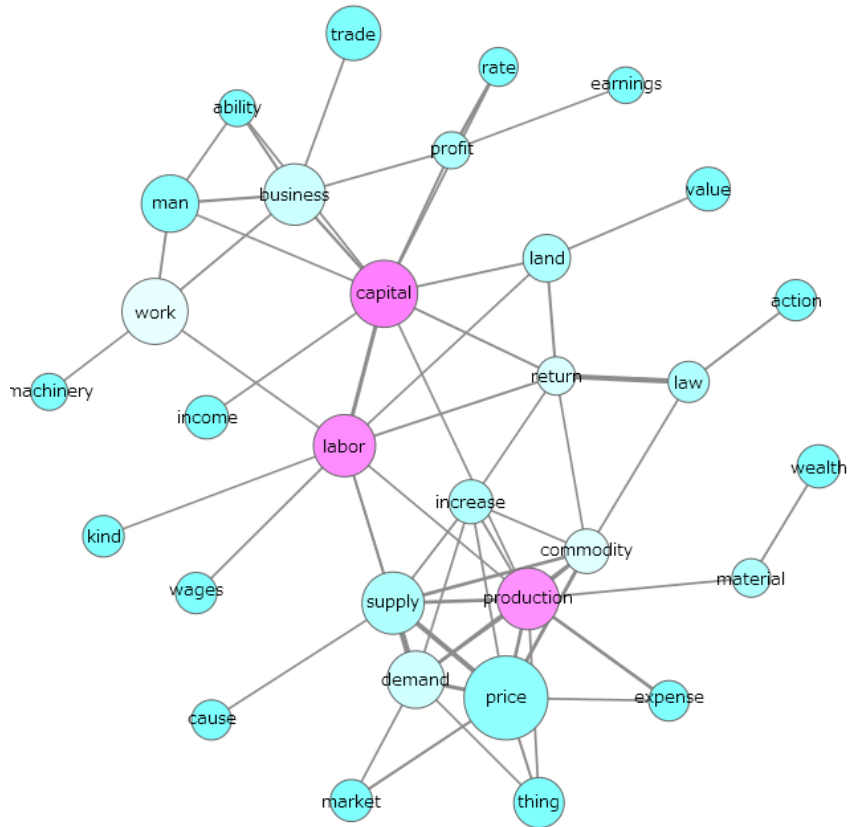
Figure 3. Network Figure of the first edition of Marshall's *Principles* (Community Analysis)
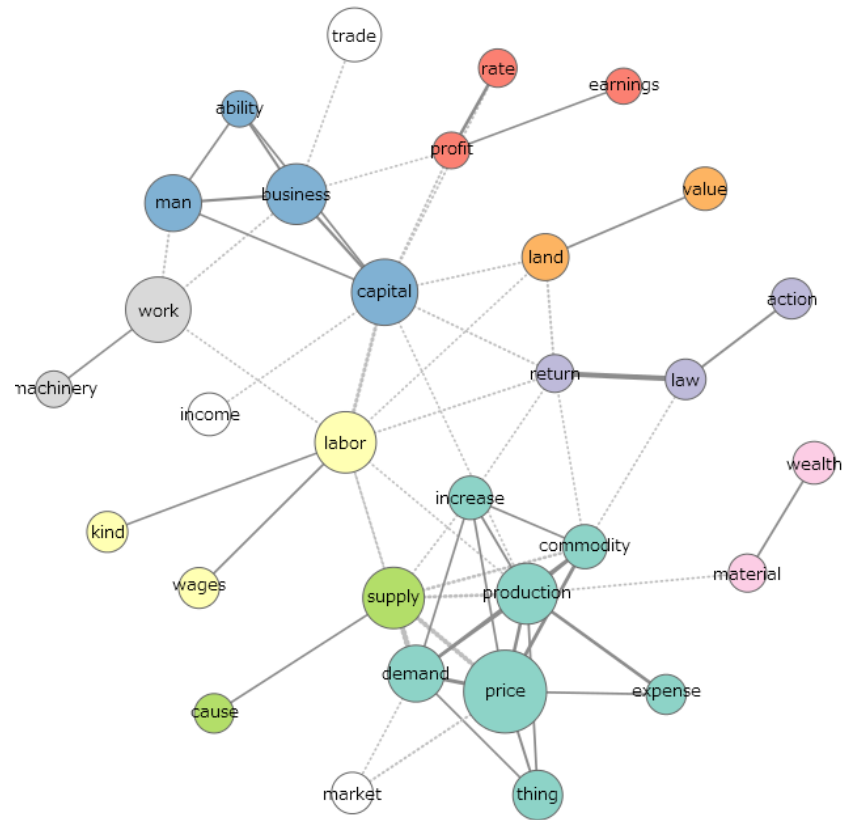
* TN: total number of words, Min TF: minimum term frequency, NoSW: number of selected words

Figure 4. Network Figure of Book II of Marshall's *Principles* (Centrality Analysis) (TN:13,662, Min TF:12, NoSW:53)*
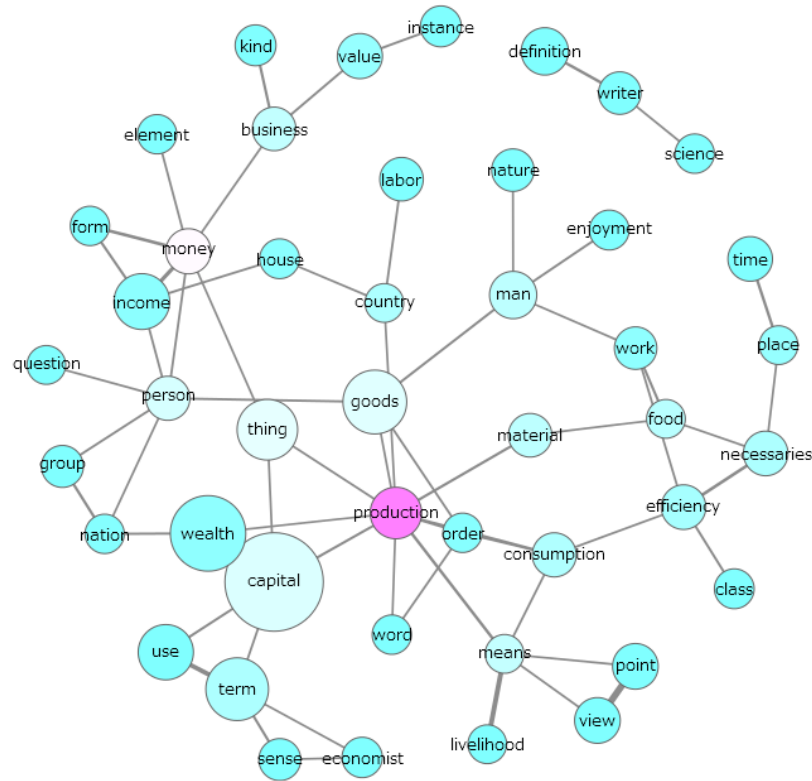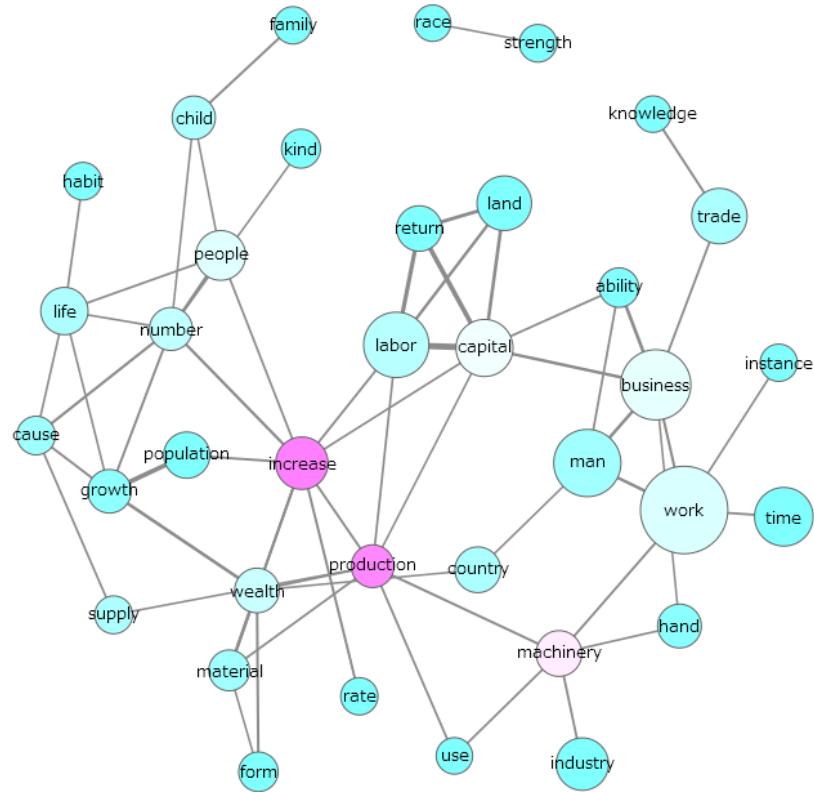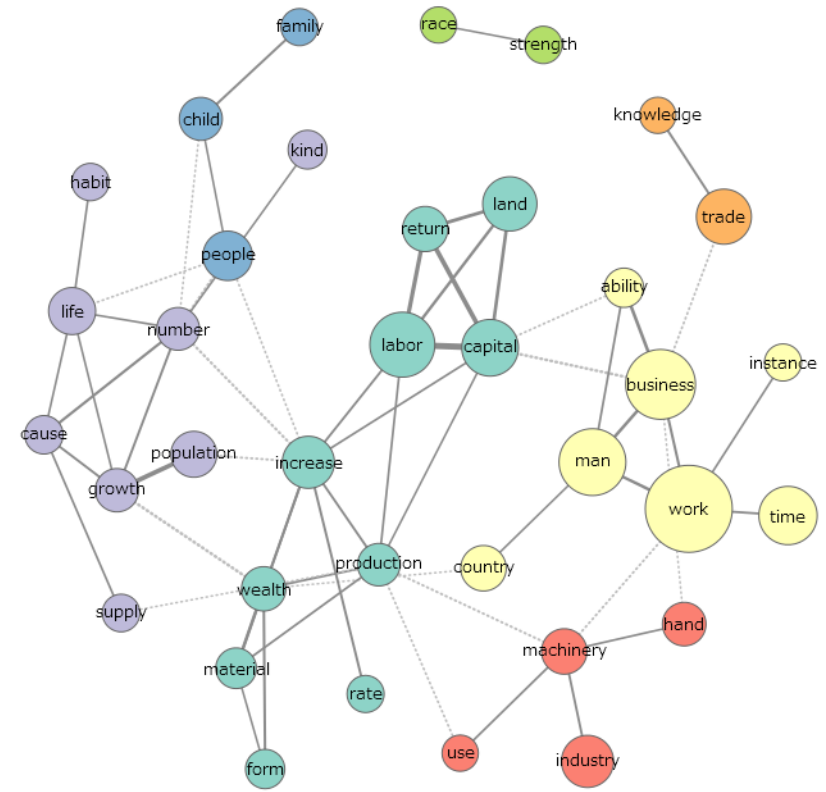

Figure 5. Network Figure of Book II of Marshall's *Principles* (Community Analysis)
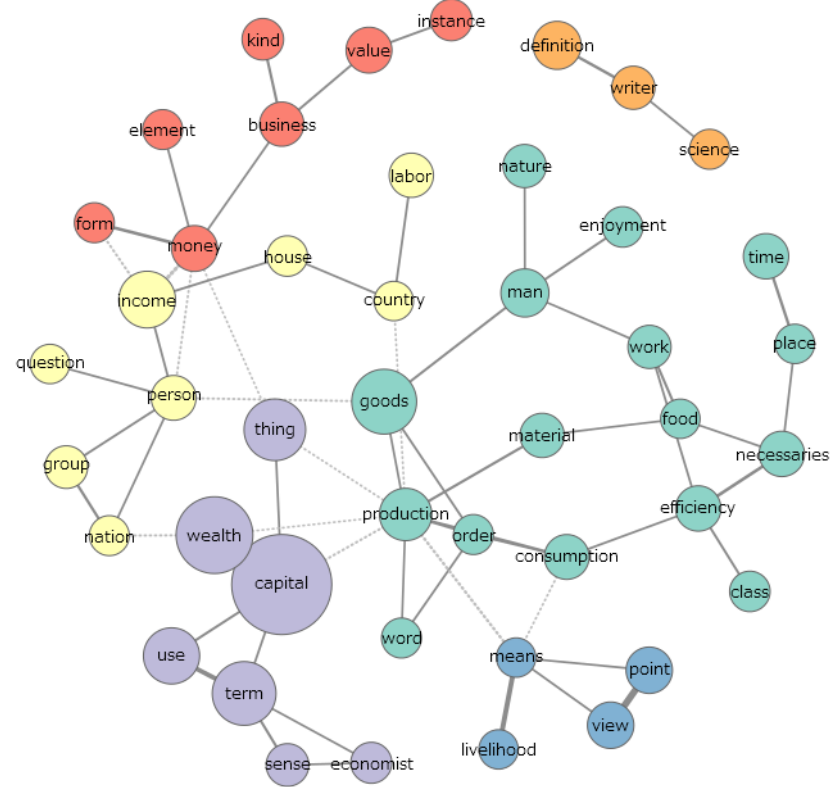
* TN: total number of words, Min TF: minimum term frequency, NoSW: number of selected words

Figure 6. Network Figure of Book IV of Marshall's *Principles* (Centrality Analysis) (TN: 64,081, Min TF:50, NoSW:48)*

Figure 7. Network Figure of Book IV of Marshall's *Principles* (community Analysis)

* TN: total number of words, Min TF: minimum term frequency, NoSW: number of selected words

Figure 8. Network Figure of Book V of Marshall's *Principles* (Centrality Analysis) (TN: 28,362, Min TF:20, NoSW:52)*
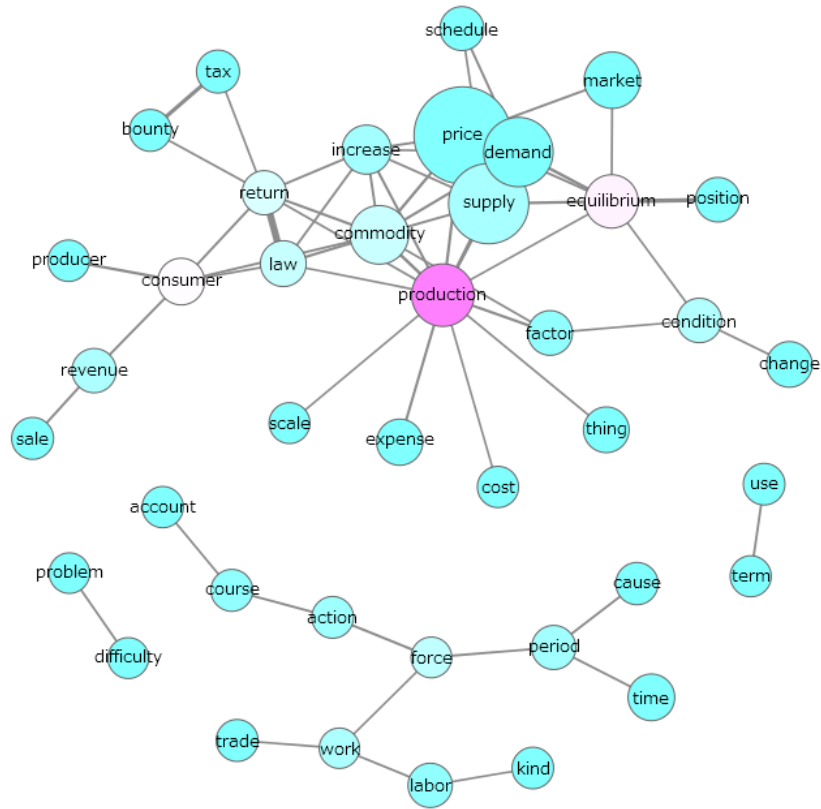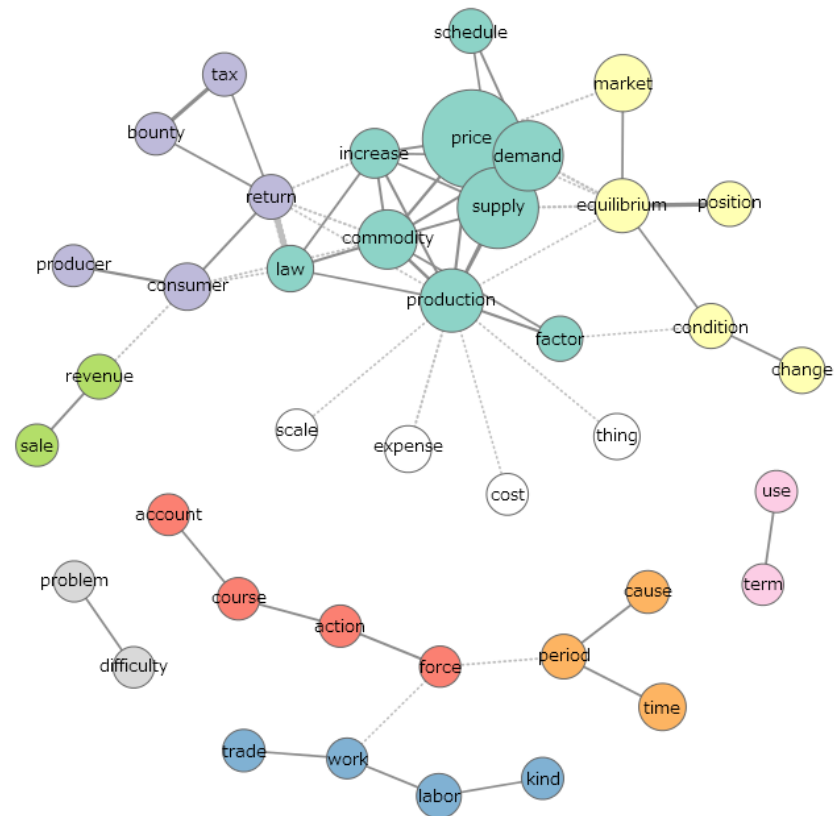
Figure 9. Network Figure of Book V of Marshall's *Principles* (Community Analysis)

* TN: total number of words, Min TF: minimum term frequency, NoSW: number of selected words

Figure 10. Network Figure of Book VI of Marshall's *Principles* (Centrality Analysis) (TN: 16,650, Min TF:15, NoSW:48)*
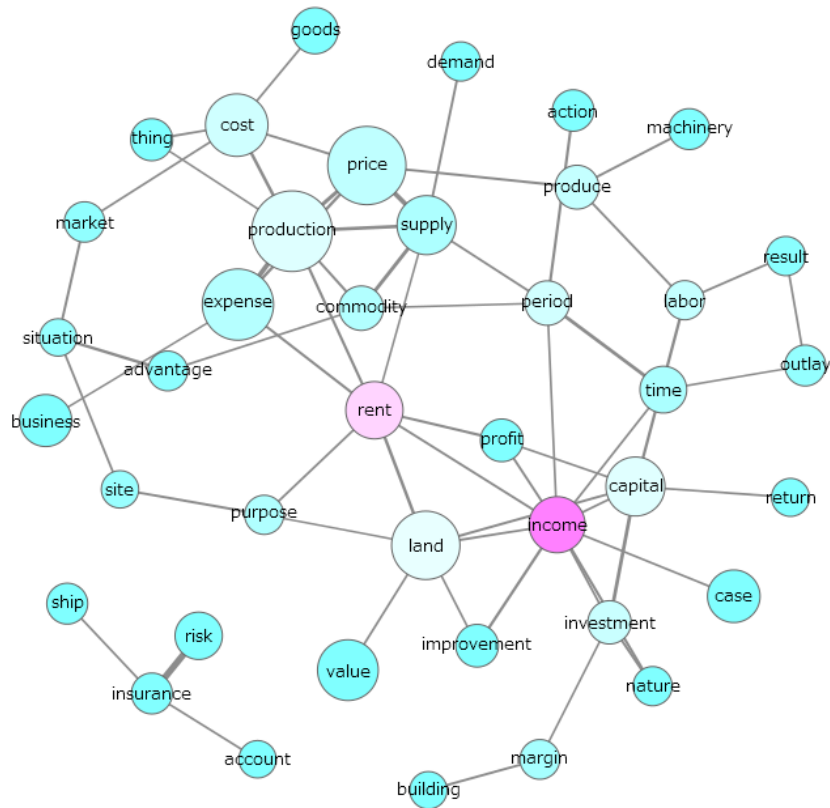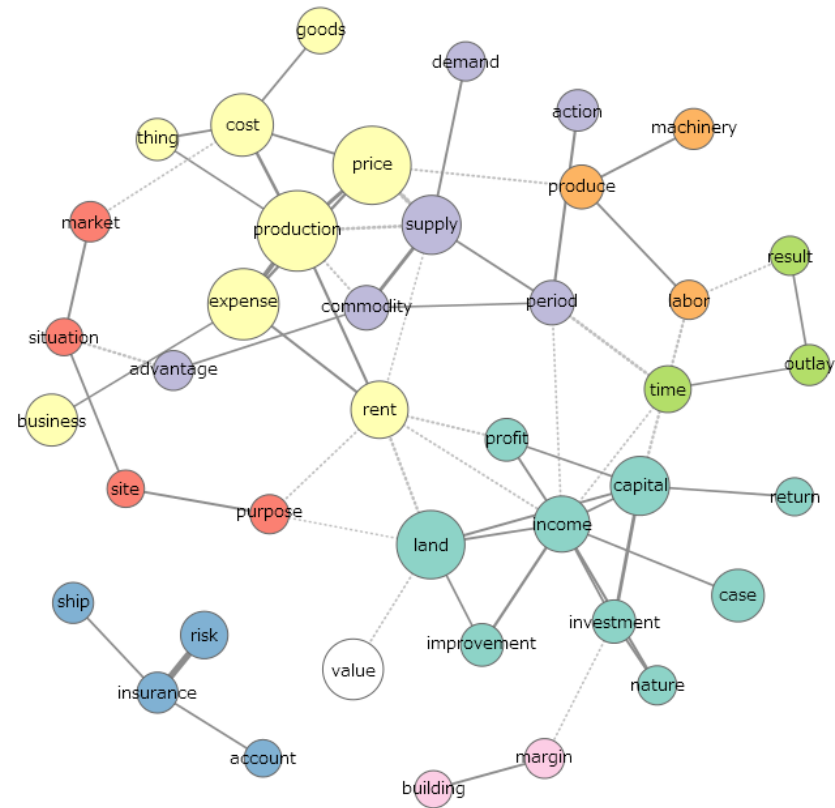

Figure 11. Network Figure of Book VI of Marshall's *Principles* (Community Analysis)

* TN: total number of words, Min TF: minimum term frequency, NoSW: number of selected words

Figure 12. Network Figure of Book VII of Marshall's *Principles* (Centrality Analysis) (TN: 60,728, Min TF:45, NoSW:50)*
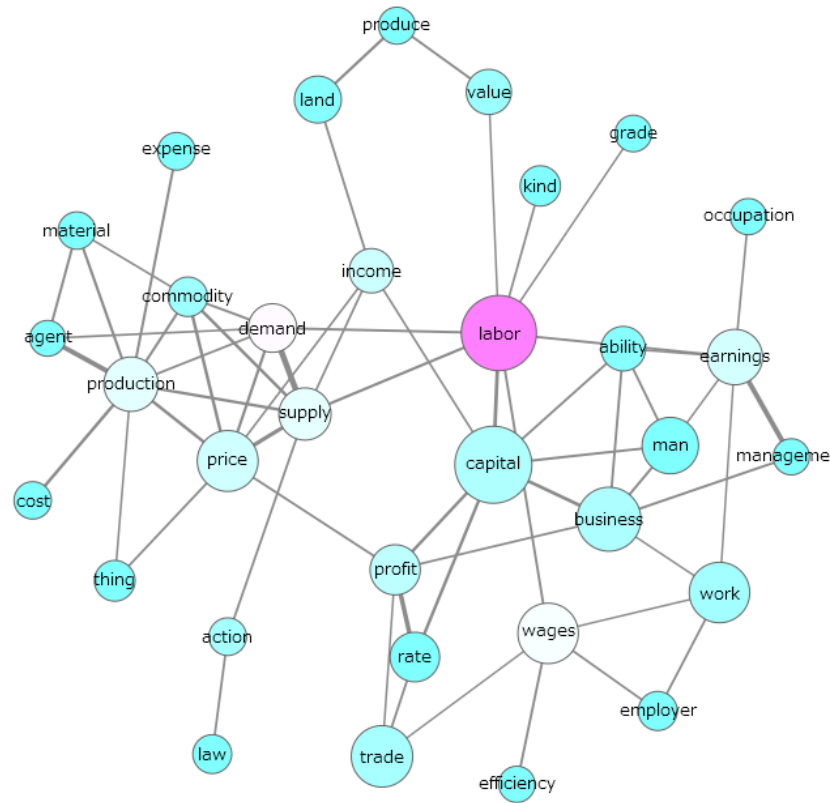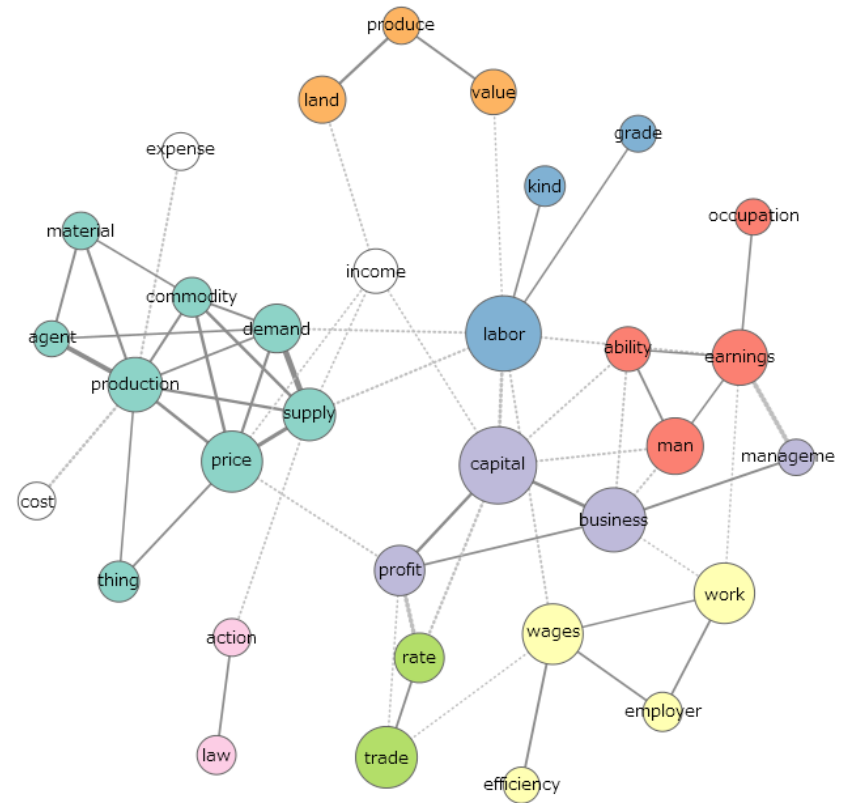
Figure 13. Network Figure of Book VII of Marshall's *Principles* (Community Analysis)

* TN: total number of words, Min TF: minimum term frequency, NoSW: number of selected words

Table 3. The Detailed Results of Text-Mining applied to Marshall's *Principles*

（1）	（2）	（3）	（4）	（5）	（6）	（7）
Book II	Some Fundamental Notions	1. production	production, order, material, goods…	Capital Community	×	capital
		2. money	money, form, element, business…		×	
Book IV	Production or Supply	1. increase	increase, production, wealth, material, labour, capital	Capital Community	×	work
		1. production				
Book V	The Theory of the Equilibrium of Demand and Supply	1. production	production, supply, price, commodity, law	Production Community	○	price
		2. equilibrium	equilibrium, market, position, condition		×	
		3. consumer	consumer, producer, return, tax, bounty		×	
Book VI	Cost of Production Further Considered	1. income	income, land, profit, capital, investment…	Production Community	×	production
		2. rent	rent, production , cost, price, business		△	
Book VII	Value, or Distribution and Exchange	1. labour	labour, kind, grade	A. Capital Community B. Labour Community	A，B=×，○	capital
		2. wages	wages, work, employer, efficiency		A，B=×，×	
		3. demand	demand, supply, price, production, commodity		A，B=×，×	

（1） The characteristic component of Marshall's *Principles*, which is based on the results of Corresponding Analysis

（2） Title of each Book of Marshall's *Principles*

（3） The central noun in each component　＊'increase' and 'production' are the same degree of centrality in Book IV.

（4） The components of each community

（5） The representative community in each Book

（6） The similarity of representative community between each Book and Marshall's *Principles*

（7） The highest frequency in use in each Book of Marshall's *Principles*