

# Big data Research on the Application of Tracking Policy Audit in the Implementation of Tax Reduction and Fee

Keqing Lei

Northwest University of Political Science and Law University, Xi'an 710122, China

**Abstract:** Since 2018, Chinese government have conducted tracking audits on the implementation of tax reduction and fee reduction policies, facing the challenges of insufficient audit resources and massive unstructured data storage. This article attempts to introduce big data technology into tracking audit policies. This article elaborates on the importance and urgency of tracking and auditing the implementation of tax and fee reduction policies. The impact of big data technology in the field of auditing on the tracking and auditing of the implementation of tax and fee reduction policies is also discussed. It is proposed that auditing agencies should apply big data technology to the tracking and auditing of the implementation of tax and fee reduction policies. And we also attempted to design an audit plan based on Neo4j graphical database technology, which includes audit objectives and scope, audit content and focus, specific big data audit steps and audit results. In response to the problems that hinder the application of big data technology, suggestions have been put forward to improve the standardization of the data process system, actively promote more efficient big data technology, and increase the training of big data audit talents.

**Keywords:** Big data auditing; Reduce taxes and fees; Policy implementation tracking audit.

## 1. Introduction

In order to better achieve tax reduction and fee reduction tracking audits, auditors need to continuously expand the sources of audit data and collect richer data to support audit conclusions. Traditional audit methods are outdated and unable to achieve a large amount of data processing. In addition to continuously improving the current policy implementation tracking audit theory, the audit department also needs to use new big data technologies to form strong support for audit work.

In the context of the big data era, computer professionals have developed numerous big data technologies to solve the storage, management, and utilization problems of big data. By using big data technology, auditors can efficiently manage and analyze audit big data, shorten the time for data management and analysis, and thus shorten the time required for audit projects. Furthermore, audit agencies can build big data platforms to collect real-time data from audited entities, quickly share audit data and findings, and promptly identify problems that arise during the implementation of tax and fee reduction policies. And the introduction of big data technology into the tracking audit of the implementation of tax and fee reduction policies not only meets the requirements of full audit coverage, but also greatly improves the quality and timeliness of audits, fully plays the supervisory role of audits, reveals the problems and shortcomings in the implementation process of tax and fee reduction policies, and helps the smooth transformation of economic structure and the achievement of high-quality economic development.

## 2. Literature Analysis

### 2.1. Research on Tax Reduction and Fee Reduction Audit

The tracking audit of the implementation of tax and fee reduction policies is divided into two sub modules: tax reduction audit and fee reduction audit. The tax reduction audit focuses on tax reduction and exemption, while the fee

reduction audit focuses on administrative violations of fees. Bi Huarong (2015) analyzed the current situation of tax reduction and refund auditing and proposed suggestions from three perspectives: institutional norms, talent development, and policy optimization. Li Ye (2017) studied the actual situation of a certain county and pointed out that there were problems such as insufficient attention and outdated technical methods in the audit of the county, and proposed corresponding solutions. Zhang Nuying (2014) summarized and summarized the audit focus of this audit project. Zhou Shuda (2014) pointed out in his research that when auditing fees for enterprises, auditing agencies should focus on the compliance of fee setting projects and the compliance and soundness of fee audit approval systems.

### 2.2. Research on the Application of Big Data Technology in Government Audit

Han Qiang (2015) believes that introducing big data technology into government auditing can effectively promote the transformation of audit analysis mode from traditional verification to mining. Wei Xiangjian (2016) pointed out that establishing a unified big data audit platform by auditing agencies can achieve. Audit resource sharing, audit strategy collaboration, and audit business collaboration. The Wuhan Special Office of the Audit Office (2018) pointed out that unstructured data in the big data environment has high information value and needs to be further explored. The constantly developing new audit technologies are powerful tools for policy implementation tracking and auditing. Liu Guocheng (2019) pointed out that in the big data environment. The audit data of policy tracking audit has the characteristics of fragmentation and massive volume, and the amount of unstructured data shows a geometric growth. Auditors need to use big data technologies such as web scraping and graphical database technology to manage and effectively analyze the data of poverty alleviation policy tracking audit. Hong Zhizhuang (2020) used the audit conducted by the Audit Bureau of H City as a research case, pointing out that auditors can effectively improve audit

efficiency and quality by introducing big data technology into policy implementation tracking audits, and assist auditors in better grasping audit priorities. The Shanghai Special Office of the Audit Office (2020) used Python technology to extract text, perform similarity and dependency calculations, and conduct data mining and correlation analysis on various types of data such as industry and commerce, finance, and taxation, improving the accuracy and efficiency of discovering audit doubts.

At present, there is sufficient literature on the application of big data technology in policy implementation tracking and auditing. However, due to the fact that tracking and auditing of tax reduction and fee reduction policy implementation is a newly launched audit business, there is still a lack of literature on how to apply big data technology to tracking and auditing of tax reduction and fee reduction policy implementation.

### **3. The Impact of Big Data Technology on Policytracking Audit**

#### **3.1. Analysis the policy tracking audit**

##### **3.1.1. Status and necessity**

Since 2016, new tax and fee reduction policies have been introduced every year. Auditors must always pay attention to the latest developments in tax and fee reduction policies in order to accurately grasp the audit objectives and content. By tracing the development process of China's tax and fee reduction policy, auditors can also have a more comprehensive and understanding of the practical background, motivation, and significance of the policy, which helps auditors to carry out targeted audit work closely around the original intention of the tax and fee reduction policy. Through tracking and auditing the implementation of tax and fee reduction policies, audit agencies can promptly identify problems and deviations in the specific process of tax and fee reduction policies, urge audited units to take timely corrective measures, and ensure that tax and fee reduction policies are effectively implemented. At the same time, by regularly issuing audit result announcements, the Audit Office can disclose the problems that exist in implementing tax reduction and fee reduction policies in various regions, playing a good role in public opinion supervision.

#### **3.2. Problems Faced**

The tracking audit of the implementation of tax and fee reduction policies can be divided into two modules: tax reduction audit and fee reduction audit. Each module involves a large number of audited objects, with tight audit tasks and a large audit workload. Audit agencies will still face many problems and challenges when conducting tax reduction and fee reduction audits, and auditors need to take measures to solve them

##### **3.2.1. Insufficient audit strength**

The Audit Office issued relevant audit guidance in 2018, which pointed out the audit focus: the first type is to reduce the tax burden on enterprises, and the second type is to standardize administrative fees. From the announcement of audit results, it can be seen that audit agencies in various regions have conducted relatively little audit work. This indicates that the audit resources and strength of audit institutions in various regions are insufficient, making it difficult to carry out audit work on all the policy covered objects mentioned above

#### **3.2.2. Excessive reliance on audited entities**

Auditors cannot obtain audit data and information without the cooperation of the audited entity. The degree of cooperation of the audited entity will affect the efficiency of auditors in collecting audit data and information. If the audited entity subjectively does not actively cooperate and delays the provision of information, auditors will need to spend a lot of time and energy communicating and negotiating with the audited entity, which will seriously slow down the progress of the audit project.

## **4. Design of Big Data Audit Plan for Tax and Fee Reduction Policies**

### **4.1. Data collection**

The data on tax and fee reduction mainly comes from the National Taxation Bureau and relevant audited units. The audit data required by auditors can be divided into two types: private data and external data. Private data refers to the financial and business data of the audited unit. Private data is difficult to obtain through public channels and requires the audited unit to actively provide it or through the dispatch of auditors to collect it on-site. External data refers to publicly available and searchable data, which can be obtained without the audited unit, such as auditors crawling the data of the required website through web crawler technology.

### **4.2. Data preprocessing**

To avoid misleading audit results due to low-quality data, auditors must perform data preprocessing on preliminary data. Data analysis requires ensuring six characteristics: data consistency, accuracy, completeness, timeliness, credibility, and interpretability. Due to the large scale and complex structure of audit big data, the audit preprocessing process takes a long time and is the most time-consuming stage in the data mining process. There are four methods for auditors to preprocess low-quality data: (1) data cleaning. The purpose of data cleaning is to remove duplicate data, irrelevant data, and inconsistent data. Data cleaning requires setting appropriate cleaning rules to avoid erroneous removal of valid data.(2) data integration. The purpose of data integration is to integrate different sources of independent data into a single logical space or physical space, thereby eliminating the drawbacks of disconnected data and connecting data islands.(3) data transformation. The purpose of data transformation is to remove irrelevant noise and normalize data by changing the representation of the original data to adapt to data mining needs.(4) data specification. The purpose of data specification is to use a small but representative new data set for data analysis while maintaining the original appearance of the data as much as possible. Auditors use the data processed by the above methods to construct a graphical database, which can effectively avoid the negative impact of low-quality audit data on audit efficiency and audit quality.

### **4.3. Build a Neo4j database**

#### **4.3.1. Neo4j database related concepts**

Neo4j graph database is currently a hot research topic in graph databases. Neo4j graph database is a high-performance NoSQL graph database that stores structured data on the network rather than in tables. It is an embedded, disk-based, fully transactional Java persistence engine, but it stores structured data on the network (called a graph from a mathematical perspective) rather than in tables. Neo4j graph

database has the following three advantages: (1) Neo4j graph database can automatically establish relationships while creating nodes, avoiding processing in complex query scenarios.(2) The underlying layer of Neo4j graph database directly stores nodes and relationships in the form of a graph, and the query time complexity remains constant.3) Neo4j graph database can visually present entities and relationships.

#### **4.3.2. Neo4j database processing advantages**

First, the data operation speed of Neo4J graph database is faster. The association between tax and fee reduction data is often hidden and complex, requiring multiple relationship points to be crossed, which requires the establishment of multiple intermediate tables, which is time-consuming, laborious, and prone to error. The data storage mode of Neo4j graph database is more efficient, without the need to establish intermediate tables, which has great speed advantages. Secondly, the Neo4j graph database stores data more flexibly and conveniently. Neo4j graph database abandons the way of storing data in the form of tables in relational databases, and chooses to store data in the form of nodes and edges. In the unstructured data storage mode, Neo4j graph database only needs to consider node attributes and edge attributes when adding new data. In the audit of the implementation of the tax reduction and fee reduction policy, auditors need to collect, store, manage, and analyze the tax reduction and fee reduction data in real time, so as to ensure the timeliness and accuracy of the audit results. From this perspective, Neo4j graph database is more suitable for big data audit of the implementation of the tax reduction and fee reduction policy than traditional relational databases.

#### **4.3.3. Graph database construction for big data audit of the implementation of tax and fee reduction policies**

Building a graph database for big data audit of tax reduction and fee reduction involves data collection, data preprocessing, entity relationship extraction, non-relational storage using Neo4j, and establishing a function module for auditing tax reduction and fee reduction. After preprocessing the audit big data through data cleaning, data protocol and other technical means, auditors can extract data entities and data relationships from the data. Then, using data entities as nodes and data relationships as edges, a Neo4j graph database can be constructed. Data nodes and data relationships are displayed in the Neo4j graph database in the form of visualization of the Neo4j graph, and auditors can intuitively and quickly use the data analysis results to discover audit suspicions.

For unstructured data types, such as text, images, and other information, extracting key content is the key step in building a graph database. For unstructured audit data, auditors can only use it to build a Neo4j graph database after extracting the required key information from it. Therefore, auditors need to extract keywords from the preprocessed unstructured audit data, and use the keywords to reflect the relationships between nodes in the Neo4j graph database. In specific operations, auditors first need to enter relevant files, obtain word vectors through word segmentation technology, eliminate semantic ambiguity information, obtain wordnet, form a keyword graph, and form a reference dictionary. Then, using a Bayesian classifier, supported by prior probabilities, and referring to the constructed dictionary, the keywords are classified to obtain structured data results.

## **4.4. Data analysis**

The auditors can use the preprocessed data for data analysis. The auditors should use big data technology as a means, audit business processes as a reference, and follow relevant audit laws and regulations to carry out big data audit work according to audit content and audit priorities. The structure of the Neo4j graph database is similar to that of social networks, and it is constructed by establishing relationships between various subjects and incidental attributes. Like knowledge graphs, it is a visualization technology in big data technology, and it can quickly find potential relationships between all subjects through query statements. When querying traditional relational databases, the interaction between tables reduces query efficiency. The non-relational database structure constructed by the Neo4j graph database can quickly find mutual relationships in a very short time, thus finding audit doubts.

## **4.5. Summarize data doubts and obtain evidence**

After finding abnormal data through data analysis, auditors will take it as an audit doubt for further investigation.

Processing. The auditors of the data analysis team will summarize the discovered audit doubts and then allocate the abnormal data to the corresponding audit evidence collection team according to the different audited units. Upon receiving the audit doubts, the audit evidence collection team shall inquire and communicate with the responsible person of the audited unit regarding these abnormal data, and request relevant information verification. After determining the audit facts, evidence collection shall be carried out to form audit evidence. It is important to note that the follow-up audit of the implementation of tax reduction and fee reduction policies is a popular project that has been gradually carried out in the past two years. The experience of each audit project of tax reduction and fee reduction is valuable and is an important reference object for similar audit projects in the future. Therefore, the big data audit process, audit indicators set and final audit drafts constructed by the auditors during the audit should be stored completely to form audit knowledge and summarize experiences and lessons.

## **4.6. Design results**

Through the above-mentioned big data audit steps, auditors not only obtain reliable audit data, but also quickly. Accurately locate audit doubts and extend the evidence collection for these audit doubts. The audit results can be obtained under the guidance of relevant regulations. The audit results of the implementation of the tax reduction and fee reduction policy should Issuing strictly in accordance with the audit content and various aspects of key concerns.

## **5. Suggestions for Tracking and Auditing the Implementation of Tax and Fee Reduction Policies Promoted by Big Data Technology**

### **5.1. Develop institutional norms for the data process**

Data is always the core of big data auditing, and when revising big data auditing standards, it is necessary to improve the data related aspects. Institutional norms for each link. In

specific big data audits, the data stage mainly involves data collection and data prediction Processing and data storage. As for the audit object, we strictly urge the audited unit to cooperate with the requirements of the audit authority and provide the necessary data and information for the audit in a timely and proactive manner. For audit units that delay providing data or intentionally fail to cooperate, relevant national departments shall promptly hold the responsible persons of the units accountable. As for the auditing entity, the auditing authority should establish standard specifications for the collection of auditing data. When formulating standards and specifications for audit data collection, audit agencies should pay special attention to the scope, content, and collection cycle of audit data. In the context of the big data era, audit agencies often perform data correlation analysis on data from multiple departments simultaneously Analysis. Audit data comes from different departments and units, and there are significant differences in the generated data. Therefore, auditors often need to spend a lot of time and effort preprocessing these data to meet the requirements of data analysis. This phenomenon cannot be solved by auditing agencies. When formulating industry data standards and specifications, relevant national departments should fully consider the requirements of audit agencies for audit data, so that the data of each department can be used more efficiently and conveniently by audit agencies. Establishing a sound and standardized system is the best measure to prevent risks. Audit agencies must develop sound data management systems and ensure that these systems are strictly implemented. The data management system should focus on the security of audit data during transfer, storage, and use. In addition, emergency handling standards should be established for possible events such as data loss and data leakage, in order to minimize the negative impact of data management errors.

## 5.2. Actively promoting the application of more efficient big data technology in auditing

The National Audit Office has always been encouraging innovative audit techniques and methods. Auditors should actively respond to the number Zhao actively explores the application of advanced and efficient big data technology tools on the basis of existing structured queries. Order. Previously, many audit agencies have actively introduced big data technology into audit projects and achieved good results. For structured data, auditors can directly access stored data through Neo4j graphical data. For unstructured data, auditors need to write database code statements, and Neo4j graphic data needs to be read and stored by running code statements. In the tracking audit of the implementation of tax and fee reduction policies, a large amount of audit data is stored in an unstructured form. Unstructured data processing is relatively complex and requires high technical capabilities. Auditors need to spend a lot of time and energy processing unstructured data. To change this situation, it is recommended that auditors actively apply more efficient big data technology tools.

## 5.3. Accelerate the cultivation of big data auditing talents

The professional competence of auditors is a key factor in determining the quality of audit projects. If there is a shortage of auditors. Without professional competence in big data auditing, the quality of big data auditing cannot be guaranteed. Big data audit review. The requirements for the professional competence of accounting personnel are reflected in two aspects: first, the ability to use big data technology, and second, the auditing industry. Ability to handle tasks. Auditors must be familiar with the audit business process in order to fully leverage the role of big data technology. By utilizing one's own big data expertise, one can quickly and efficiently connect the relationships between different parts of data.

## References

- [1] Huarong Bi, Strengthening the Audit of Tax Reduction and Exemption in Local Tax Audit, Chinese and Foreign Entrepreneurs. 25(2015), p.109+111.
- [2] Ye Li: Research on Discipline Inspection and Supervision in Chinese Towns and Towns. (master, Jilin University, China 2017).
- [3] Zhizhuang Hong: Research on Application Big Data Technology in Policy Tracking Audit (master, Zhejiang University of Business and Technology, China 2020)
- [4] Jun Chen, Research on Innovative Audit Techniques and Methods Driven by Full Audit Coverage, Audit Research. 5(2018), p.22-29.
- [5] Nuying zhang, Suggestions Strengthening Grassroots Audit Work, Shanxi Finance and Taxation, 07(2014) No. 425, p.35-36.
- [6] Shuda Zhou, Strengthening Audit Supervision to Promote Reducing the Burden on Enterprises, China Small and Medium Enterprises, 11(2015) No.240, p.27.
- [7] Research Group of Wuhan Special Mission Office, Research on Tracking Audit Innovation and Development of Implementation of Major National Policy Measures, Audit Research, 04(2018), p.18-23.
- [8] Research Group of the Shanghai Special Mission Office, Tracking and Review of the Implementation of Big Data Technology in Major National Policy Measures, Application Research in Accounting, Audit Research, 02(2020), P.14-21.
- [9] Guocheng Liu, Huang Kun, Research on the Tracking Audit Mechanism of Poverty Alleviation Policies, Audit Research, 03(2019), No.209, p.11-19.
- [10] Alex Schara schicin, Tom McBnde, Policy analytics and accountability mechani sms: judging the value for money of policy mplementation, Annals of Operations Research, 01(2016), No.236, p.33-36.
- [11] Gisléia Benini Duarte, André de Souza Melo, Diego Firmino Costa da Silva. Do government audits reduce dengue? Estimating the impact of federal monitoring lotteries program on dengue incidence. International Journal of Health Economics and Management, (3-4)2019, No.19, p.45-48.
- [12] Garry Kim. The Application of Artificial Intelligence in Auditing: Looking Back to the Future. Expert Systems with Application, 2012, p.90-95.