The governance role of big data tax enforcement: evidence from related party transaction

Liguang Zhang[1]

zlg1990zlg@163.com

College of Management, Sichuan Agricultural University

No.211, Huimin Road, Wenjiang District, Chengdu city, Sichuan Province, China

Lisha Tan

2013852563@qq.com

College of Management, Sichuan Agricultural University

No.211, Huimin Road, Wenjiang District, Chengdu city, Sichuan Province, China

Yunchen Wang

wangyc8684@qq.com

College of Management, Sichuan Agricultural University

No.211, Huimin Road, Wenjiang District, Chengdu city, Sichuan Province, China

Liao Peng

pengliao_93@163.com

School of Accountancy, Southwestern University of Finance and Economics

No. 55 Guanghuacun Street, Chengdu city, Sichuan Province, China

[1] Please address correspondence to Liguang Zhang via email: zlg1990zlg@163.com

Acknowledgements

Liguang Zhang gratefully acknowledges financial support from the System Science and Enterprise Development Research Center of Sichuan, China (Grant Number: Xq22C15) and Chengdu Philosophy and Social Science Planning Project in 2022(Grant Number: 2022BS100). Yunchen Wang gratefully acknowledges financial support from the National Social Science Fund of China (Grant Number: 18CGL016).

The governance role of big data tax enforcement: evidence from related party transaction

Abstract: Related party transactions (RPTs) are the main way of tunneling for controlling shareholders. As big data tech becomes prevalent, we investigate how the big data in tax enforcement affects RPTs. Employing a quasi-natural experiment, we hypothesis and find evidence that the big data in tax enforcement significantly curbs firms' RPTs. The results remain stable under a series of robustness check. We further find that big data restrict RPTs by reducing the information asymmetry between tax authorities and firms and improving the tax enforcement ability of authorities. The association is stronger in firms suffering from more serious tunneling, worse corporate governance, greater tax avoidance motivation, in more tax enforcement incentive regions, and in more technological developed regions. Moreover, we provide evidence that big data in tax enforcement can significantly improve the firm value of those with a large proportion of RPTs, indicating a corporate governance role of big data tech. This study supplements the empirical evidence for whether RPTs can create firm value and provides certain policy reference significance for government to develop big data technology and strengthen the supervision of RPTs.

Key words: Big data; Tax enforcement; RPTs; Corporate governance

1. Introduction

The rapid development of information technology represented by big data, cloud computing, artificial intelligence, and their extensive use in economic and social life have greatly improved the decision-making efficiency of economic bodies. Digital information has become an indispensable factor of production, as important as traditional factors such as labor, capital, or land. The way to seize the historic opportunity of the new industrial revolution to improve the integrated development

of big data and the real economy has also been urgent to find out. In the process of deep integration in big data technology and traditional way of life and work, it is necessary to consider how big data application affects the operational efficiency of economic bodies and its underlying mechanism. The answers will lay the foundation for vigorously developing big data technology and improving the efficiency of economic operation.

This paper examines how the application of big data in tax enforcement plays a role in governance. The government's ability to collect taxes and fiscal capacity is one of the important cornerstones of national prosperity (Besley and Persson, 2012). How to improve the efficiency of tax enforcement has received widespread attention from the government and academia. However, limited by human, material, and financial resources, governments are often unable to effectively combat corporate tax evasion. In 2013, China's government started to introduce a new tax enforcement information system, namely the third stage of the China Tax Administration Information System (CTAIS-3) pilot, which applies emerging information technologies such as big data, cloud computing, and artificial intelligence to tax enforcement. It can obtain tax related information about taxpayers' investment, production, distribution and consumption, greatly strengthen the ability of tax related information analysis and supervision, and significantly crack down on corporate tax evasion and tax evasion, improving the efficiency of tax law enforcement. It marks that China's tax enforcement has entered the era of big data collection and management from manual collection and management (Xiao and Shao, 2020; Li et al., 2020).

Related-party transactions (RPTs, hereafter) are defined as transactions with related entities such as shareholders, members of the board of directors, and affiliated companies. Although reasonable RPTs can reduce transaction costs within group companies, improve the efficiency of capital utilization, and achieve economies of scale (Cook, 1977; Fishman and Khanna, 2004), RPTs in China often become a major mean for controlling shareholders or management (i.e., insiders) hollow out listed companies (Cheung et al., 2006; Jian and Wong, 2010; Berkman et al., 2010). Since

tunneling listed companies will reduce corporate taxable profits and thus corporate tax revenues, we expect that big data tax enforcement will strengthen the supervision of corporate RPTs behaviors, reducing the RPTs behaviors of tunneling and playing a governance role.

Taking the CTAIS-3 as instance, we believe that the impact of big data tax enforcement on RPTs is mainly reflected in two aspects: First, big data tax enforcement reduces the information asymmetry between tax authorities and enterprises. The CTAIS-3 covers all tax types, all taxation work process, all state and local taxation agencies as well as relevant departments such as industry and commerce, statistics and banks, etc. Tax authorities can obtain tax-related information of RPTs from different sources, including taxpayers' tax declaration information, taxpayers' information reported to banks and statistical departments and information collected on the Internet, so as to cross-verify data from enterprise tax declaration with those from other departments. These data basically cover all processes of the enterprise's production and operation, that is, fully reflect the actual production and operation status of the enterprise, alleviating the information asymmetry between tax authorities and corporations, helping the tax authorities to strengthen the supervision of tax-related information of enterprise RPTs, and finally curbing tax avoidance or RPTs with hollow motives. Secondly, big data tax enforcement improves the taxation authorities' ability to analyze and supervise tax-related information. With the help of big data technology, the CTAIS-3 monitors all the tax-RPTs of all enterprises. Moreover, the decision support system in the CTAIS-3 carries out tax big data analysis and provides early warning under comprehensive tax evaluation by industry, tax type, and region. It monitors the tax-related information of enterprises in real time and finds out suspicious information, facilitating to lock the audit objects and greatly improving the efficiency of tax enforcement. Therefore, big data tax enforcement relies much less on the number, tax expertise, and work experience of tax staff, which exactly makes up for the shortcomings of traditional tax enforcement and improves the efficiency of tax enforcement. Consequently, this study argues that big data tax enforcement can

significantly suppress the RPTs of enterprises.

Employing the CTAIS-3's implementation as a quasi-natural experiment and Chinese A-shares from 2008 to 2016 as samples, this study constructs a difference-in-differences (DID) model to examine how big data tax enforcement affects enterprise RPTs. The findings are as follows: (1) The CTAIS-3 significantly suppresses the scale of enterprise RPTs, and the result remains robust under various checks including parallel trend test, propensity score matching test and placebo test and so on; (2) The mechanism analysis indicates that the CTAIS-3 mainly works by reducing the information asymmetry between tax authorities and enterprises and improving the tax capacity of tax authorities; (3) Heterogeneity analysis finds that the effect of big data tax enforcement on RPTs is influenced by enterprises' RPTs motives, corporate governance, and regional tax motives. Specifically, the relationship between the big data tax enforcement and RPTs is more prominent in enterprises with more controlling shareholders' tunneling, with worse corporate governance, with stronger tax avoidance motives, and in regions with stronger tax motives and better information technology of tax officers; (4) By further examining enterprises' motives for RPTs, we find that big data tax enforcement can significantly increase the firm value of enterprises with a larger proportion of RPTs, thus confirming that the tunneling motivation plays a leading role in the motivation of RPTs in Chinese publicly listed companies.

Compared with prior studies, this study contributes in several ways. First, it enriches and expands the literature on economic consequences of big data information technology. Most previous studies have focused on the effects of traditional information technology on firms' own operations, such as improving the timeliness of corporate financial reporting disclosures (Brazel and Dang, 2008), reducing internal control deficiencies (Masli et al., 2010), reducing audit risk and improving audit quality (Pincus et al., 2017). Some studies have examined how big data applications affect corporate tax avoidance (Li et al., 2020; Xiao and Shao, 2020), earnings informativeness (Zhao, 2021), and investment efficiency (Zhang et al., 2022), but few

of them focused the economic consequences of the government's adoption of big data information technology from the perspective of corporate governance. Our study examines the governance role of big data tax enforcement from the perspective of RPTs to fill the gap.

Second, this study supplements the literature related to tax enforcement and microenterprise behavior from the perspective of RPTs. Existing studies have focused on how tax enforcement affects corporate tax avoidance (Hoopes et al., 2012; Lennox et al., 2022), corporate disclosure (Hanlon et al., 2014; Bauer et al., 2021), financing costs (Guedhami and Pittman, 2008), and corporate value (Dyck and Zingales, 2004; Desai et al., 2007), with less literature examining the specific ways in which tax enforcement inhibits corporate tax avoidance and plays a governance role. RPTs is a primary tool of corporate tax avoidance and tunneling by controlling shareholders (Cheung et al., 2006; Lo et al., 2010). We find that big data tax enforcement significantly reduce corporate RPTs, especially in firms with more tunneling, suggesting that discouraging RPTs may be a specific way for tax enforcement to play a positive role. It is useful for extending our understanding of the tax enforcement's role.

Third, this study expands the relevant research on the influencing factors of RPTs from the perspective of big data information technology. Existing studies have focused on how equity structure (Kang et al., 2014), director characteristics (Hope et al., 2019), auditor reputation (Bennouri et al., 2011; Fang et al., 2017), product market competition (Chen et al., 2012), short selling mechanism (Jiang et al., 2021), and other internal and external corporate governance mechanisms affect RPTs, while few studies investigated the government adoption of big data information technology in monitoring RPTs. The rapid development and widespread use of big data has profoundly changed people's life, work, and way of thinking, and is conducive to improving the efficiency of decision-making by economic agents, which also includes government departments. Examining how the tax enforcement of big data affects the RPTs behavior of enterprises can help us understand how modern information technology shapes the behavior of micro-enterprises and macroeconomic operations

and provide empirical evidence for the government to accelerate the development of big data.

Fourth, the findings further provides evidence on whether RPTs can increase enterprise value, as a policy reference for the government to strengthen the regulation of RPTs. Existing studies show that RPTs can not only achieve economies of scale by reducing transaction costs within the group (Cook, 1977; Fishman and Khanna, 2004), but also increase corporate value by reducing the overall tax burden of the group through transfer pricing (Lo et al., 2010). On the other hand, agency problems give insiders (i.e., controlling shareholders, management) an incentive to pursue self-interests through RPTs, thereby damaging corporate value (Cheung et al., 2006; Berkman et al., 2010). By using a quasi-natural experiment, the implementation of the CTAIS-3, we find that big data in tax enforcement significantly reduces corporate RPTs and increases the firm value of those with large RPTs, suggesting that the decline in RPTs facilitates to increase firm value and confirming that RPTs of Chinese listed companies are detrimental to firm value.

2. Literature review, theoretical analysis and research hypothesis

2.1. Institutional background

As an emerging market, China has a low efficiency of law enforcement with prevalent enterprise tax evasions (Cai and Liu, 2009). In order to improve the tax enforcement efficiency and reduce tax loss, the State Taxation Administration of China has continuously promoted the informatization level of tax enforcement, and the "China Tax Administration Information System" is the most representative one. At the first stage in 1994, the State Taxation Administration launched the first phase of China Tax Administration Information System (CTAIS-1) to effectively manage VAT invoices, which collected taxpayers' invoice information manually and then compared it with the computer to detect problems. However, the error rate of manual data collection was so high that the CTAIS-1 was stopped at the end of 1996. At the second stage in 2001, the second phase of the China Tax Administration Information System (CTAIS-2) was launched. It further improved the management of VAT invoices and

played a certain role in combating tax evasion and increasing tax revenue. but the CTAIS-2 has deficiencies such as only being able to manage VAT information while unable to cover the local tax system, and lack of the sharing of tax data among regions. At the third stage, the third phase of the China Tax Administration Information System (CTAIS-3) was launched in 2013 and gradually implemented to the whole country.

Compared with the CTAIS-1 and CTAIS-2, CTAIS-3 has following advantages. Firstly, CTAIS-3 covers all tax types, all tax process, state and local tax authorities and relevant departments such as industry and commerce, statistics, banks, etc. The tax authorities obtain tax-related information of RPTs from different sources, including tax declared by taxpayers, information reported by taxpayers to banks and statistics departments, and the Internet, so as to cross-check the data of enterprise tax declaration with those of other departments. These data basically cover all aspects of enterprises' production and operation, fully reflecting the actual production and operation conditions of enterprises, alleviating the information asymmetry between taxation authorities and tax-related subjects. It facilitates tax authorities to strengthen the supervision of tax-related information of enterprises. Secondly, the CTAIS-3 realizes centralized storage, centralized processing and centralized management of taxation information in the General Administration and Provincial Administration. It enables taxation authorities in different provinces to verify the cross-regional business activities and tax payment of enterprises and strengthen the sharing and exchange of taxation information among taxation authorities, hence inhibiting the tax avoidance of enterprises. Finally, the traditional tax enforcement mainly adopted the empirical management mode of "people control people" to selectively collect and investigate taxes. In contrast, the CTAIS-3 monitors all the tax-RPTs of all enterprises with the help of big data technology and uses the decision support system in the CTAIS-3 to carry out tax analysis and comprehensive tax analysis by industry, tax type and region. The big data analysis and comprehensive taxation early warning evaluation of taxation can monitor the tax-related information of enterprises in real time, detect suspicious information, and carry out early warning, which is convenient for the

taxation authorities to lock the audit objects and greatly improve the efficiency of tax enforcement. Therefore, big data tax enforcement relies much less on the number, tax expertise, and work experience of tax staff, which makes up for the shortcomings of traditional tax enforcement and improves the tax enforcement efficiency. Therefore, this paper argues that big data tax enforcement can significantly suppress enterprise RPTs.

2.2. Literature Review

2.2.1. Impact of tax enforcement on enterprise behavior

Existing studies mainly focus on the direct and indirect effects of tax enforcement on enterprises.

The direct impact of tax enforcement. Granted and stipulated by law, tax authorities embrace the power and the obligation of tax inspection. They have not only the motive but also the ability to inspect the accounting books and statements of companies to combat tax evasion by RPTs or transfer pricing. Using the probability of tax audit to measure the intensity of tax enforcement, scholars provide empirical evidence that enhanced tax enforcement discourages corporate tax avoidance. Hoopes et al. (2012) examine the relationship between tax enforcement and corporate tax avoidance in the context of the U.S. They find that corporate tax avoidance decreases significantly when the probability of Internal Revenue Service (IRS) audits increases. Using proprietary data obtained from local tax bureaus in China, Lennox et al. (2022) find that tax audits significantly increase the effective corporate tax rate and reduce book-tax differences.

Indirect effects of tax enforcement. Taxation is the government's mandatory share of corporate profits, that is, in effect the government becomes the "minority shareholder" of all firms and has an incentive to actively monitor insider misappropriation, which objectively plays a corporate governance role (Dyck and Zingales, 2004; Desai et al., 2007). Basing on data on control transactions in 39 countries, Dyck and Zingales (2004) find that the private benefits of controlling shareholders are lower in countries with more stringent tax enforcement. Desai et al. (2007) first incorporate tax enforcement into the corporate governance system using a theoretical model. They find that

increasing tax enforcement is beneficial for increasing firm value.

As taxable profits are highly consistent with accounting profits, Hanlon et al. (2014) find that strict enforcement by tax authorities reduces management's distortion of the firm's true (taxable) profits, thereby improving the quality of financial reporting. Bauer et al. (2021) further find that strict tax enforcement constrains managers' bad news hoarding activities evident in stock price crash risk, improving corporate information transparency. Correspondingly, Guedhami and Pittman (2008) find that strict tax enforcement reduces the cost of equity capital for firms by mitigating information asymmetries between firms and outside investors and reducing the risk premium demanded by investors.

2.2.2. Economic consequences of information technology

Existing literature focuses on the impact of firms' application of traditional information technology on their own operations. For example, Hitt et al. (2002) and Anderson et al. (2006) found that increased investments in information technology increase firm value; Brazel and Dang(2008) found that firms' adoption of Enterprise Resource Planning (ERP) software systems can improve the efficiency of accounting processing and increase the timeliness of corporate financial reporting disclosures, yet leading to more earnings management; Masli et al. (2010) argue that the adoption of internal control system technology can reduce internal control deficiencies and corporate information risk; Pincus et al. (2017) found that clients adopted Enterprise Systems (ESs) to increase operational transparency, improve the internal information environment, and reduce audit risk, therefore improving audit quality and audit efficiency.

With the development of modern information technology such as big data, cloud computing, and the Internet, studies have begun to focus on the impact of big data applications on corporate behavior. Zhu (2019) found that big data collected by third-party companies for listed companies such as consumer transactions and satellite images can improve the information content of stock prices and investment efficiency. Using a quasi-natural experiment implemented by China's implementation of the

CTAIS-3, Li et al. (2020) and Xiao and Shao (2020) find that the application of big data technology to tax enforcement alleviates the information asymmetry between tax authorities and enterprises, improves the capacity and efficiency of tax enforcement, and inhibits corporate tax avoidance behavior. Basing on this, Zhao (2021) finds that the CTAIS-3 improves the quality of profits, thus improving the earning informativeness. Zhang (2022) finds that the CTAIS-3 alleviates information asymmetry and improve the efficiency of enterprises' investment by improving the quality of financial reports.

Although existing studies examine the effects of tax enforcement on corporate tax avoidance and information disclosure, it is unclear in what ways tax enforcement inhibit corporate tax avoidance and play a governance role. RPTs is not only a major tool for large shareholders to transfer resources and encroach on the interests of minority shareholders (Cheung et al., 2006), but also an important instrument for group enterprises to reduce the overall tax burden (Lo et al., 2010). Therefore, this study examines the impact of big data tax enforcement on RPTs not only to identify how tax enforcement can combat corporate tax evasion, but also to understand the ways in which tax enforcement can play a role in external corporate governance.

2.3. Theoretical analysis and research hypothesis

Enterprises in most countries are controlled by an ultimate controlling shareholder (La Porta et al., 1999; Shleifer and Vishny, 1997). The controlling shareholder of the enterprise at the upper-level controls its subsidiaries by holding their controlling shares, and finally controls the listed company at the end, known as the pyramid structure. The ultimate controlling shareholder constitutes an enterprise group together with many enterprises controlled by them through pyramid structure. In emerging markets, enterprise groups as an alternative to underdeveloped external capital markets, can form content capital markets, thus easing financing constraints (Almeida and Wolfenzon, 2006). However, the increased separation of voting right and cash flow right brought about by the pyramid structure gives controlling shareholders an incentive to infringe on the interests of minority shareholders for

private gain, in which RPTs is an important tool (Cheung et al., 2006; Jian and Wong, 2010).

Tax authorities are committed to supervising taxpayers to pay taxes to ensure the source of state revenue. Given the applicable tax rate of enterprises remains unchanged, the amount of tax paid by enterprises directly depends on their business results. Therefore, tax authorities have the motivation and ability to monitor and combat any actions that would undermine the tax base. As some RPTs divert resources from listed companies and decrease the tax paid, which will damage the performance of tax enforcement agencies, tax authorities will strengthen the supervision of these two types of RPTs. After the launch of CTAIS-3, taxation authorities can use big data technology to comprehensively recognize the operations of enterprises and tax-related information in commodity transactions, accurately identifying the risk of tax evasion through the decision support system in CTAIS-3 and effectively suppressing the RPTs with tunneling motives. Specifically, big data tax enforcement can inhibit RPTs through the following two ways.

First, big data tax enforcement can reduce the information asymmetry between tax authorities and enterprises. The CTAIS-3, which integrates the information technologies such as big data, cloud computing and the Internet, covers all tax types and various departments, improving the transparency of tax-related information of enterprises' RPTs and reducing the information asymmetry, mainly in the following ways. On the one hand, tax-related information from different sources can corroborate each other. For the purpose of tunneling, insiders often construct complex and hidden RPTs, which increases the difficulty for tax authorities to identify tax-related risks. Moreover, in the past, enterprise data from various government departments were isolated from each other and information sharing and interoperability could not be achieved, giving enterprises an incentive to report different financial data to various departments in their own interests. Due to the limitation of the number of tax officers, tax authorities cannot check all the RPTs of enterprises and effectively restrain the RPTs with tunneling motive. The CTAIS-3 makes it possible for taxation authorities to

obtain tax-related information from different sources, including tax information declared by taxpayers, information reported by taxpayers to banks and statistics departments as well as information collected on the Internet. It is convenient for taxation authorities to cross-check the tax declaration data of enterprises with the data of other departments. These data fully reflect the actual production and operation status of enterprises, which can reduce the information asymmetry between taxation authorities and tax subjects and help taxation authorities to strengthen the supervision of enterprises' RPTs. For example, tax authorities can obtain information of registration and changes of enterprises from the industry and commerce department and identify related parties of listed companies through equity penetration, thereby increasing the difficulty for public companies to conceal RPTs. On the other hand, tax-related information from different regions can be shared. Before the implementation of CTAIS-3, while each province had its own tax enforcement system, the tax enforcement systems of different provinces were not connected. It is impossible for information sharing and exchange. It was costly for local authorities to obtain tax-related information of enterprises in other provinces. Since the applicable tax rates and regulatory environment vary in provinces, enterprises can easily manipulate the amount of transactions with other provinces' partners by taking advantage of the information asymmetry across regions, especially through RPTs (Xiao and Shao, 2020). As a unified tax enforcement system at the national level, the CTAIS-3 is managed by the State Taxation Administration, which enables tax authorities in different provinces to verify enterprises' cross-regional business activities and tax payments and enhances tax information sharing between tax authorities, thus discouraging enterprises' RPTs. For example, when a public company conducts RPTs with affiliated enterprises in other provinces, the tax authorities can simultaneously compare the information of the RPTs with other non-RPTs of the public company and non-RPTs of affiliated enterprises in other places, so as to check whether there are signs of tunneling.

Secondly, big data tax enforcement improves the tax ability of authorities. Traditional tax enforcement mainly adopts the empirical management model of

"people managing people" to collect and investigate taxes, which not only has a single source of information, but also relies on the experience and expertise of tax staff. In the meantime, in order to minimize the risk of detection by tax authorities, tax avoidance activities are often characterized by complexity, ambiguity and opacity (Desai and Dharmapala, 2006). Insiders also construct complex RPTs for tunneling, which makes it greatly difficult for tax officials. As the number of corporate taxpayers far exceeds the number of tax officers, it is impossible for tax officers to examine all corporate RPTs. Therefore, by the traditional method, the taxation authorities have limited tax enforcement capacity and can hardly effectively suppress the RPTs of enterprises. In contrast, the CTAIS-3 has established a decision support system for taxation economic analysis, monitoring and prediction for taxation authorities at all levels. For example, the decision support system includes six early warning indicators for enterprise tax assessment: warning of inconsistency between VAT revenue and enterprise income tax revenue, warning of abnormal inventory tax deduction ratio, warning of abnormal elasticity of changes in input tax and output tax, warning of excessive proportion of pre-received accounts, warning of abnormal income tax contribution rate, and warning of abnormal change rate of main business income. With the help of decision support system, tax authorities can easily check the income, expense, cost, profit rate, bank account, etc. related to RPTs of enterprises. By comparing the profit rate of RPTs with that of peers, officials detect suspicious information and make early warning, that is, easily lock the target of audit and greatly improve the tax enforcement efficiency. Hence, big data tax enforcement relies much less on the number, expertise, and experience of tax staff, which makes up for the shortcomings of traditional tax enforcement.

In summary, big data tax enforcement can suppress enterprise RPTs by alleviating information asymmetry and improving tax enforcement capacity. Therefore, we put forward the following hypothesis:

H1: Under other conditions unchanged, big data tax enforcement is conducive to suppressing corporate RPTs.

3. Research design, sample selection and descriptive statistics

3.1. Definition of variables

1. Related party transactions (RPTs)

This study uses RPTs to measure the tunneling behavior of large shareholders. Drawing on Jian and Wong (2010) and Lo et al.(2010), we define RPTs as transactions of goods and services between affiliated companies for following reasons. First, although RPTs of listed companies in China cover purchase or sale of assets, goods and services, equity and related guarantee, the RPTs of goods and services are the most common. Second, unlike fixed assets or equity, transactions in goods and services are recurring activities so that the tunneling activities through them is less likely to be detected (Jian and Wong, 2010). Therefore, it is difficult for the tax authorities to supervise tunneling RPTs with limited human and material resources. After the CTAIS-3 is launched, the taxation authorities can use big data technology to inspect the business conditions of enterprises and tax-related information in commodity transactions, accurately identifying the risk of tax evasion by enterprises through the decision support system in the CTAIS-3 and effectively suppressing the tunneling behavior. C Consequently, big data tax enforcement has a more direct impact on the RPTs of goods and services.

2. Big data tax enforcement

Following Xiao and Shao (2020) and Li et al. (2020), this study measures big data tax enforcement (*GTP*) using the implementation of the CTAIS-3. If a firm is located in a region where the CTAIS-3 has been implemented in the current year, the *GTP* variable takes the value of 1, and 0 otherwise.

3.2. Model setting

Drawing on Beck et al. (2010), this study constructs a staggered difference-in-difference (DID) model and controls for both firm (*Firm*) and year (*Year*) fixed effects in the model to mitigate the potential bias from omitted variables at the firm level. We construct the following regression model:

$$RPT = \beta_0 + \beta_1 GTP + \sum \beta_i Controls + \sum Firm + \sum Year + \varepsilon \quad (1)$$

RPT is an indicator for RPTs, which is equal to the total amount of goods and services transactions occurring between listed companies and related parties divided by the total assets by the end of the year. *GTP* is an indicator for CTAIS-3, which takes the value of 1 if a firm is located in a region where the CTAIS-3 has been implemented in the current year and 0 otherwise. Following existing studies, *Controls* denotes a series of control variables, including corporate financial characteristics such as company size, financial leverage, and growth, and corporate governance characteristics such as ownership concentration and board independence. Details of variable definition are shown in Table 1. Moreover, model (1) incorporates regional GDP growth rate and tax growth rate to control for the effects of local economic development and tax growth. To control for the effects of serial autocorrelation, we adjust the regression residuals for clustering at the firm level. The interest of variable is the coefficient β1 in model (1), which measures the net effect of big data tax enforcement on large shareholder tunneling.

Table 1 Definition of variables

Variable Symbol	Variable Definition Description
RPT	RPTs, which equals to the total amount of goods and services transactions occurring between listed companies and related parties divided by total assets at year-end.
GTP	The *GTP* dummy variable, which takes the value of 1 if a firm is located in a region where the CTAIS-3 has been implemented in the current year and 0 otherwise.
Size	Firm size, which equals to the natural logarithm of the firm's total year-end market capitalization.
Lev	Financial leverage, which equals to the company's total year-end liabilities divided by total year-end assets.
Growth	Firm growth, which equals to (current year sales revenue - prior year sales revenue) / prior year sales revenue.
Fcf	Free cash flow, which equals to the difference between cash flow from operating activities and capital expenditures divided by total assets at year-end.
Top1	Concentration of shareholding, which equals to the number of shares held by the largest shareholder divided by the total share capital at the end of the year.
Top1_25	Shareholding balance, which equals to the shareholding of the first

	largest shareholder divided by the sum of the shareholding of the second to fifth largest shareholders.
Boardsize	Board size which equals to the natural logarithm of the board members at the end of the year.
Indep	Board Independence, which equals to the number of independent directors divided by the number of directors at the end of the year.
Ins_ratio	Institutional ownership, which equals to the number of shares held by institutional investors divided by the number of floating stock shares at the end of the year.
Taxrate	Nominal income tax rate, which equals to the statutory corporate income tax rate.
AbsDA	Earning management, which equals to the absolute value of the residuals obtained by regressing the extended Dechow and Dichev (2002) model by year and by industry.
G_growth	Regional GDP growth rate, which equals to (current year GDP - previous year GDP) / previous year GDP.
Tax_growth	Regional tax revenue growth rate, which equals to (total tax revenue of the current year - total tax revenue of the previous year / total tax revenue of the previous year.

3.3. Sample selection and data sources

The sample period of this paper is during 2008-2016 for several reasons: (1) the new enterprise income tax regulation, which started to be implemented on January 1, 2008, unifies the income tax rate of domestic and foreign enterprises at 25%. To prevent the impact of the new enterprise income tax law on enterprises' motive of tunneling, the sample period of our study starts from 2008; (2) by the end of 2016, all provinces in China have implemented the CTAIS-3. As we take the provinces where the CTAIS-3 was implemented in the second half of the year as the second year of implementation, the sample period of this paper ends in 2016 to ensure that a control group sample always exists during the sample period, consistent with the approach of Zhang et al. (2020). Consequently, this study selects all A-share listed companies from 2008-2016 as the initial sample and screens the research sample as follows: (1) remove companies in financial industry; (2) remove companies in abnormal operation such as ST and PT; (3) remove companies with asset-liability ratio greater than 1 and other variables with missing values. In addition, the extreme values of all continuous variables are winsorized at 1%. After the above processing, we finally obtain 13,821 firm-year observations in our sample. The data of financial characteristics are obtained from

CSMAR database. The data of the implementation time of CTAIS-3 in each province are manually compiled from the official website of the State Taxation Administration and news reports. The data of the statutory corporate income tax rate is obtained from Wind database.

3.4. Descriptive statistics

Table 2 presents the descriptive statistics of the main variables. The mean and standard deviation of RPT are 0.042 and 0.092, respectively, indicating that RPTs vary significantly among different public companies. The mean value of *GTP* is 0.113, indicating that 11.3% of the observations are in the years after the implementation of the CTAIS-3. The mean and median of *G_growth* were 0.095 and 0.087, indicating that the economic development of provinces and cities is quite fast during the sample period. The results of descriptive statistics for other control variables are similar to the existing literature.

Table 2 Descriptive statistics

Variables	Observation	Mean	SD	1/4 quartile	Median	3/4 quartile
RPT	13821	0.042	0.092	0.000	0.005	0.036
GTP	13821	0.113	0.316	0.000	0.000	0.000
Size	13821	22.616	0.944	21.900	22.500	23.200
Lev	13821	0.436	0.209	0.267	0.434	0.602
Growth	13821	0.209	0.441	-0.003	0.129	0.301
Fcf	13821	-0.004	0.116	-0.038	0.016	0.055
Top1	13821	0.365	0.153	0.241	0.349	0.473
Top1_25	13821	4.634	6.775	1.070	2.060	4.880
Boardsize	13821	2.274	0.178	2.200	2.300	2.300
Indep	13821	0.372	0.053	0.333	0.333	0.400
Ins_ratio	13821	6.416	5.774	2.030	4.760	9.020
Taxrate	13821	0.187	0.054	0.150	0.150	0.250
AbsDA	13821	0.086	0.086	0.027	0.060	0.113
G_growth	13821	0.095	0.023	0.078	0.087	0.109
Tax_growth	13821	0.114	0.067	0.082	0.109	0.154

4. Empirical Analysis

4.1. Correlation analysis

Table 3 shows the result of correlation analysis for the main variables of model (1). There is a significant negative correlation between the *GTP* and the *RPT*, initially

verifying that the tax enforcement of big data plays a governance role and inhibits the tunneling of large shareholders, supporting our hypothesis. The correlation coefficients between the control variables are mostly below 0.3, indicating that model (1) does not have serious multicollinearity problems. In order to mitigate the effects of other factors on the relationship, we conduct multiple regression analysis below.

Table 3 Table of correlation coefficients for each variable

	RPT	*GTP*	*Size*	*Lev*	*Growth*	*Fcf*	*Top1*	*Top1_25*	*Boardsize*	*Indep*	*Ins_ratio*	*Taxrate*	*AbsDA*	*G_growth*	*Tax_growth*
RPT	1.000														
GTP	-0.011**	1.000													
Size	0.124***	0.150***	1.000												
Lev	0.180***	-0.017**	0.157***	1.000											
Growth	-0.008	-0.009	0.099***	0.050***	1.000										
Fcf	0.039***	-0.009	-0.004	0.043***	-0.097***	1.000									
Top1	0.217***	-0.054***	0.193***	0.112***	0.003	0.069***	1.000								
Top1_25	0.258***	-0.039***	0.041***	0.171***	-0.069***	0.062***	0.536***	1.000							
Boardsize	0.182***	-0.030***	0.146***	0.175***	-0.024***	0.013	0.030***	0.049***	1.000						

Indep	-0.047***	0.007	0.081***	-0.011	0.016**	-0.002	0.063***	0.023***	-0.456***	1.000					
Ins_ratio	-0.024***	0.037***	0.287***	0.022**	0.094***	-0.061***	-0.188***	-0.191***	0.012	-0.008	1.000				
Taxrate	0.071***	-0.003	0.142***	0.311***	0.001	0.018**	0.128***	0.149***	0.055***	0.023***	-0.051***	1.000			
AbsDA	-0.046***	0.026***	0.037***	-0.025***	0.254***	-0.221***	-0.008	-0.047***	-0.080***	0.020**	0.046***	0.001	1.000		
G_growth	0.114***	-0.230***	-0.254***	0.098***	0.041***	-0.029***	0.032***	0.079***	0.114***	-0.043***	-0.014	0.005	-0.031***	1.000	
Tax_growth	0.054***	-0.337***	-0.223***	0.050***	0.051***	0.006	0.024***	0.069***	0.073***	-0.019**	-0.047***	0.000	-0.047***	0.664***	1.000

Note: *, ** and *** denote 10%, 5% and 1% significance levels, respectively.

4.2. Big data tax enforcement and RPTs: multiple regression analysis

Table 4 reports the results of multiple regression of big data tax enforcement and RPTs. The dependent variable is RPTs (*RPT*) and the independent variable is the implementation of CTAIS-3 (*GTP*). In column (1), in order to investigate the direct effect of the CTAIS-3 on RPTs, only firm-year fixed effects are controlled in the model. The coefficient of *GTP* is significantly negative at the 1% statistical level, indicating that the RPTs of enterprises is significantly reduced after the implementation of the CTAIS-3. In columns (2)-(3), we successively add the variables of enterprise financial characteristics, corporate governance characteristics and macroeconomic characteristics to the model, still finding that the coefficient of *GTP* is still significant at the 1% level. It indicates that big data tax enforcement can significantly reduce the scale of corporate RPTs and inhibit the tunneling by controllers. In terms of economic significance, after the implementation of the CTAIS-3, the RPTs of enterprises decreased by 14.29%[]. The stepwise regression results in Table 4 robustly confirm our hypothesis that big data tax enforcement reduces corporate RPTs, inhibits the tunneling by major shareholders, and exerts governance effects.

Table 4: Big data tax enforcement and RPTs

	(1) RPT	(2) RPT	(3) RPT
GTP	-0.007***	-0.006***	-0.006***
	(-2.93)	(-2.94)	(-2.88)
Size		0.003	0.003
		(1.52)	(1.48)
Lev		0.007	0.007
		(0.93)	(0.91)
Growth		0.004***	0.004***
		(2.67)	(2.66)
Fcf		0.004	0.004
		(1.01)	(0.98)
Top1		0.025	0.025
		(1.41)	(1.41)
Top1_25		0.001***	0.001***

[] 14.29% is equal to the regression coefficient of GTP -0.006 in column 3 divided by the mean value of RPT 0.042 and taken as the absolute value.

		(3.24)	(3.24)
Boardsize		0.017*	0.017*
		(1.78)	(1.79)
Indep		0.026	0.026
		(1.13)	(1.14)
Ins_ratio		0.000	0.000
		(0.31)	(0.32)
Taxrate		0.026	0.025
		(0.86)	(0.85)
AbsDA		-0.003	-0.003
		(-0.49)	(-0.50)
G_growth			0.005
			(0.07)
Tax_growth			0.007
			(0.49)
Constant	0.041***	-0.104*	-0.104*
	(31.19)	(-1.78)	(-1.76)
Firm	Yes	Yes	Yes
Year	Yes	Yes	Yes
N	13821	13821	13821
Adj.R2	0.01	0.02	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by company-level clustering (Cluster).

4.3. Big data tax enforcement and RPTs: robustness tests

4.3.1. Parallel trend test and propensity score matching test

The parallel trend assumption is an important prerequisite for the adoption of the DID model. Drawing on Bertrand and Mullainathan (2003), this study tests the parallel trend assumption by examining the dynamic impact of the CTAIS-3 on RPTs. Specifically, we add time dummy variables *Before1*, *After0*, *After1*, and *After2_* into model (1), which take the value of 1 before 1 year, the current year, the next 1 year, the next 2 years and subsequent years of the implementation of the CTAIS-3 respectively, otherwise 0. Column (1) of Table 5 illustrates that the coefficient of *Before1* is not statistically different from 0, which indicates no statistically difference in the scale of RPTs between the treatment and control group enterprises 1 year before the implementation of the CTAIS-3. In contrast, the regression coefficients of *After1* and

*After2_* are both significantly negative at the 5% statistical level, indicating that the RPTs' scale of the treatment group enterprises declines significantly more than that of the control group enterprises. These results indicate that there are parallel trends in the changes of RPTs between the treatment and control group firms 1 year before the implementation of the CTAIS-3, and these trends show significant differences only after the implementation of the CTAIS-3. The results support that the our sample satisfies the parallel trend hypothesis, confirming the validity of the DID model design.

To further mitigate the interference of the differences in firm characteristics between the treatment and control groups on the study conclusions, we employ the propensity score matching (PSM) method to construct paired samples to repeat the regression analysis. Specifically, following DeFond et al. (2015), we add all control variables in model (1) to the logit regression to ensure that all the characteristic factors that may affect RPTs are not significantly different between the treatment and control groups. Then, basing on the calculated propensity scores, we use the caliper matching technique (without replacement) to match the treatment and the control group firms 1:1, thus obtaining the paired sample. The regression results are shown in column (2) of Table 5. It shows that, consistent with the full-sample regression results, the regression coefficient of *GTP* is still significantly negative at the 5% statistical level, further demonstrating that big data tax enforcement can inhibit large shareholders from tunneling, which is consistent with the findings in Table 4.

Table 5 Big data tax enforcement and RPTs: parallel trend test and PSM test

	(1) RPT	(2) RPT
Before1	0.000	
	(0.16)	
After0	-0.002	
	(-0.98)	
After1	-0.008**	
	(-2.31)	
*After2_*	-0.013***	
	(-3.13)	
GTP		-0.009**
		(-2.43)

Size	0.003**	0.005
	(2.42)	(1.40)
Lev	0.007	-0.016
	(1.39)	(-1.20)
Growth	0.004***	0.004
	(3.46)	(1.63)
Fcf	0.004	0.003
	(1.04)	(0.29)
Top1	0.026***	0.001
	(2.96)	(0.02)
Top1_25	0.001***	0.001***
	(6.64)	(3.80)
Boardsize	0.017***	0.024
	(3.10)	(1.58)
Indep	0.027*	0.019
	(1.82)	(0.47)
Ins_ratio	0.000	0.000
	(0.51)	(0.28)
Taxrate	0.025	0.062
	(1.39)	(1.13)
AbsDA	-0.003	-0.017
	(-0.54)	(-1.46)
G_growth	0.007	-0.810**
	(0.14)	(-2.41)
Tax_growth	0.007	0.060*
	(0.62)	(1.87)
Constant	-0.096***	-0.052
	(-3.00)	(-0.50)
Firm	Yes	Yes
Year	Yes	Yes
N	13821	2704
Adj.R2	0.02	0.08

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors were adjusted for firm-level clustering (Cluster).

4.3.2. Placebo test

Following Fauver et al. (2017), in Table 6 we conduct a placebo test to enhance the reliability of the findings from baseline. Specifically, we only retain the sample before the implementation of the CTAIS-3 and then advance the implementation of the CTAIS-3 in each province by 3 years and 5 years, respectively, as the pseudo

implementation year of the CTAIS-3 implementation in each province. We generate the pseudo CTAIS-3 variables *Pseudo_GTP3* and *Pseudo_GTP5,* and repeats the regression of model (1). The results are presented in column (1) - (2). It can be found that the regression coefficients of *Pseudo_GTP3* and *Pseudo_GTP5* are statistically insignificant.

Table 6 Big data tax enforcement and RPTs: placebo test

	(1)	(2)	(3)
	RPT	RPT	RPT
Pseudo_GTP3	0.001		
	(0.69)		
Pseudo_GTP5		0.002	
		(0.93)	
Random_GTP			0.002
			(1.01)
Size	0.002	0.002	0.004
	(0.74)	(0.73)	(1.58)
Lev	0.007	0.007	0.007
	(0.92)	(0.93)	(0.94)
Growth	0.004**	0.004**	0.004***
	(2.50)	(2.50)	(2.63)
Fcf	0.006	0.006	0.004
	(1.34)	(1.34)	(1.00)
Top1	0.027	0.027	0.024
	(1.35)	(1.35)	(1.38)
Top1_25	0.001**	0.001**	0.001***
	(2.54)	(2.54)	(3.22)
Boardsize	0.012	0.012	0.017*
	(1.25)	(1.26)	(1.78)
Indep	0.025	0.025	0.027
	(0.98)	(0.98)	(1.15)
Ins_ratio	0.000	0.000	0.000
	(0.39)	(0.39)	(0.29)
Taxrate	0.019	0.019	0.025
	(0.57)	(0.56)	(0.82)
AbsDA	-0.001	-0.001	-0.003
	(-0.25)	(-0.24)	(-0.48)
G_growth	-0.012	-0.025	0.022
	(-0.17)	(-0.34)	(0.32)
Tax_growth	0.003	0.004	0.012
	(0.17)	(0.26)	(0.79)
Constant	-0.056	-0.055	-0.115*

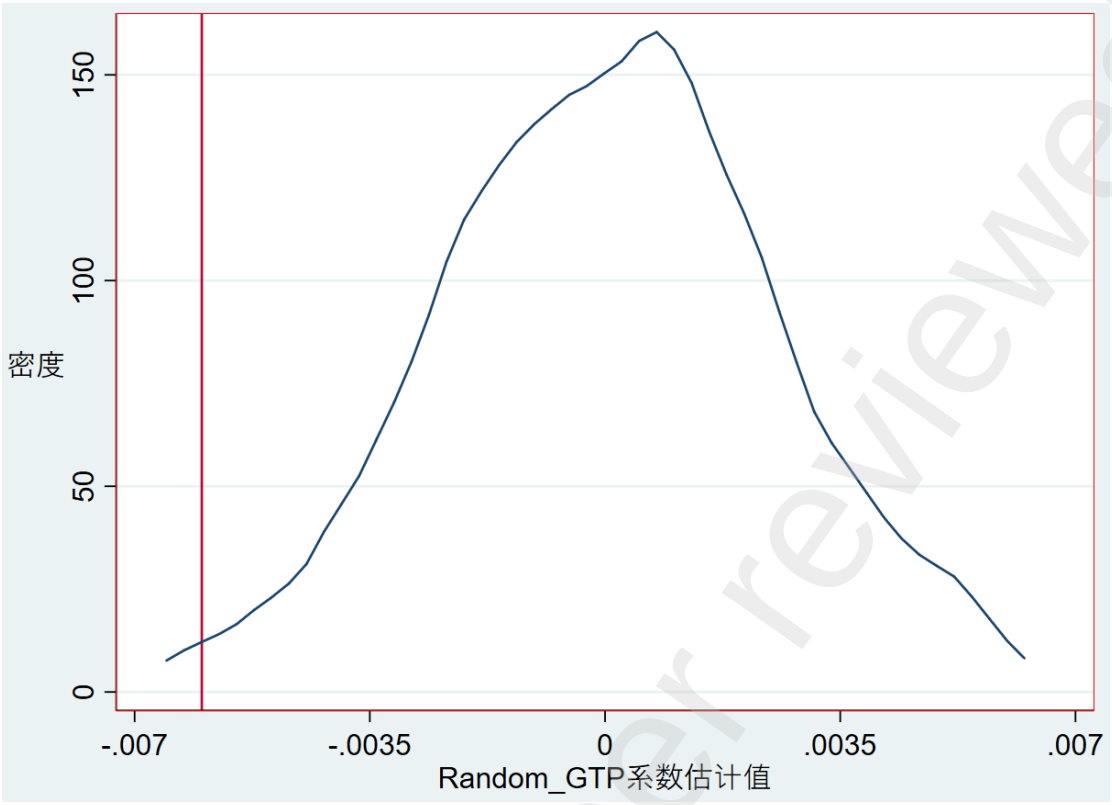
	(-0.88)	(-0.87)	(-1.91)
Firm	Yes	Yes	Yes
Year	Yes	Yes	Yes
N	12261	12261	13821
Adj.R2	0.01	0.01	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by firm-level clustering (Cluster).

To further exclude the influence of unobservable factors under the findings, drawing on Liu and Lu (2015), we conduct a placebo test. We randomly assign the year of implementation of the CTAIS-3 to each province, then generating the pseudo CTAIS-3 variable *Random_GTP* basing on random years. After rerunning the model (1), the results are presented in column (3) of Table 6. If the decline in the size of RPTs is caused by other unobservable random factors during the sample period, the regression coefficient of *Random_GTP* should still be significantly negative. However, as can be seen from column (3) of Table 6, the regression coefficient of *Random_GTP* is positive but insignificant, suggesting that the findings of this paper are not brought by unobservable random factors but by the implementation of CTAIS-3.

In addition, we repeated the above random process 500 times and saved the estimated regression coefficients. The regression coefficients distribution of the 500 randomly generated *Random_GTP* is shown in Figure 1. Consistent with our expectations, the distribution of the regression coefficients is centered at 0 (mean -0.00004). Moreover, the coefficient estimated in the baseline is -0.006 (i.e., -0.006 as indicated by the vertical line in Figure 1), exceeding the 1% quantile of the 500 random coefficients distribution (i.e., the 1% quantile is -0.005). The results in column (3) of Table 6 and Figure 1 indicate that the conclusions of this study are not subject to interference from unobservable factors.

The results of the placebo test imply that the pseudo CTAIS-3 did not lead to differences between the treatment and control groups in terms of RPTs, which further supports the conclusion that the differences in the trends of RPTs' scale between the treatment and control groups are brought by the implementation of the real CTAIS-3.



Figures 1. The density distribution of the estimated coefficients from 500 simulations using pseudo date of the CTAIS-3 pilot.

4.3.3. Sample Replacement

In recent years, the China's government has carried out a series of reforms on corporate taxation, among which the most influential one is the reform of business tax replaced with value-added tax (BT-VAT reform). Moreover, the original purpose of the "BT-VAT reform" is to strengthen tax enforcement and to combat tax evasion by enterprises. In order to exclude the interference of the " BT-VAT reform" policy with the basic conclusions of this paper, we remove the samples of the real estate, transportation, storage and postal, and construction industries that levied business tax before the BT-VAT reform and repeat the regression of model (1). The regression results are shown in column (1) of Table 7. It can be found that after excluding the sample of " BT-VAT reform", the coefficient of GTP is still significantly negative at the 1% statistical level, which is consistent with the findings in Table 4.

In order to more accurately capture the impact of the CTAIS-3's implementation on analysts' forecasts and reduce the interference of other factors caused by the long

research period to the conclusions, we limit the research period to 2 years before and after the implementation of the CTAIS-3 to repeat the regression of model (1). The regression results are shown in column (2) of Table 7. Meanwhile, in order to cover as many samples as possible and improve the representativeness of research conclusions, we extend the research period to 2008-2019. The regression results are shown in column (3) of Table 7. It can be found that in columns (2)-(3), the coefficients of the *GTP* are still significant at the 1% statistical level, indicating that the conclusion of baseline is still valid.

Table 7 Big data tax enforcement and RPTs: replacement sample

	(1) Delete the "BT-VAT reform" sample RPT	(2) Shorten the period [-2, +2] RPT	(3) Set the period to 2008-2019 RPT
GTP	-0.006***	-0.005***	-0.005***
	(-2.98)	(-2.90)	(-3.37)
Size	0.005***	0.007***	0.002
	(2.84)	(2.66)	(1.14)
Lev	0.010	0.002	0.004
	(1.60)	(0.29)	(0.67)
Growth	0.006***	0.002	0.004***
	(3.79)	(1.60)	(3.99)
Fcf	0.005	-0.002	0.004
	(1.06)	(-0.32)	(1.22)
Top1	0.041***	0.028*	0.029**
	(3.73)	(1.72)	(2.01)
Top1_25	0.001***	0.001**	0.001***
	(5.82)	(2.14)	(3.00)
Boardsize	0.018**	0.011	0.010
	(2.54)	(1.06)	(1.27)
Indep	0.043**	0.017	0.023
	(2.34)	(0.83)	(1.16)
Ins_ratio	0.000	-0.000	0.000
	(1.51)	(-0.59)	(0.74)
Taxrate	0.061***	-0.006	0.027
	(2.70)	(-0.13)	(1.00)
AbsDA	-0.002	-0.009	0.001
	(-0.34)	(-1.37)	(0.28)
G_growth	-0.002	0.406**	0.065
	(-0.03)	(2.41)	(1.19)

Tax_growth	0.009	-0.023	0.005
	(0.65)	(-1.15)	(0.43)
Constant	-0.149***	-0.193***	-0.051
	(-3.63)	(-2.97)	(-1.13)
Firm	Yes	Yes	Yes
Year	Yes	Yes	Yes
N	9959	6958	19662
Adj.R2	0.03	0.02	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by firm-level clustering (Cluster).

4.3.4. Other robustness tests

The government may adopt different tax incentives and regulatory policies for different industries, thus affecting the behavior of enterprises in conducting RPTs. The interaction term between industry dummy variables and annual dummy variables is added to model (1) to capture the impact of industry differences on the findings. The regression results are shown in column (1) of Table 8. It can be found that the regression coefficients of *GTP* are still significantly negative, and the findings remain unchanged.

The measurement of RPTs may also have potential interference on the baseline results. We use the following two ways to re-measure RPTs: (1) *RPT_relative*, which equals to the total amount of RPTs occurring between a listed company and its parent company, other subsidiaries in the conglomerates, and companies controlled by its top management divided by the total assets by the end of the period; (2) *Adj_RPT* is the RPT adjusted according to the industry-annual median, which equals to the RPT of the company in the current year minus the median of the industries' RPT in the current year. The regression results are shown in columns (2)-(3) of Table 8. It can be found that the regression coefficient of GTP is still significantly negative even if the new measurement method of RPTs is adopted and the significance level and the magnitude of the regression coefficient do not change substantially, indicating that the conclusion of our baseline is hardly affected by the measurement method of RPTs .

Since tax expenses are an important part of corporate cash outflows, big data tax

enforcement may have a negative impact on corporate operating performance by compressing the space for corporate tax avoidance. Therefore, the decrease in the scale of RPTs may also be caused by the CTAIS-3 reducing the growth rate of enterprises. In order to exclude this alternative explanation, we adopt the growth rate of business income (*Salsgrowth*) and total assets (*Assetgrowth*) to measure enterprise growth, further testing whether the implementation of the CTAIS-3 damages enterprise growth. The regression results are shown in columns (4)-(5) of Table 8. It can be found that the regression coefficients of GTP are insignificantly negative, implying that the implementation of CTAIS-3 does not damage enterprise growth on the whole. As such, we exclude the potential explanation that the conclusion of baseline result is caused by the decline of enterprise growth.

Table 8 Big data tax enforcement and RPTs: Others robustness tests

	Higher order fixed effects of industry and year	Changing the measurement of RPTs		Exclusion of alternative explanations	
	(1)	(2)	(3)	(4)	(5)
	RPT	RPT_relative	Adj_RPT	Salsgrowth	Assetgrowth
GTP	-0.005***	-0.006***	-0.005***	0.007	-0.007
	(-2.85)	(-3.04)	(-2.85)	(0.41)	(-0.50)
Size	0.002*	0.002	0.002	0.264***	0.186***
	(1.82)	(0.89)	(1.44)	(17.47)	(16.17)
Lev	0.009*	0.005	0.007	0.597***	0.449***
	(1.91)	(0.89)	(1.46)	(10.67)	(8.80)
Growth	0.003***	0.003**	0.003***		
	(2.80)	(2.41)	(3.26)		
Fcf	0.003	-0.000	0.003	-0.039	-1.058***
	(0.81)	(-0.03)	(0.79)	(-0.83)	(-21.89)
Top1	0.026***	0.021	0.024***	0.611***	-0.045
	(2.97)	(1.40)	(2.84)	(5.12)	(-0.52)
Top1_25	0.001***	0.000**	0.001***	-0.007***	-0.001
	(6.30)	(2.32)	(6.25)	(-6.01)	(-1.48)
Boardsize	0.014***	0.015*	0.015***	0.040	0.038
	(2.58)	(1.86)	(2.85)	(0.73)	(0.89)
Indep	0.024	0.021	0.023	0.182	-0.003
	(1.62)	(1.12)	(1.59)	(1.25)	(-0.03)

Ins_ratio	0.000	0.000	0.000	-0.002	0.000
	(0.83)	(0.42)	(0.75)	(-1.39)	(0.41)
Taxrate	0.016	0.028	0.012	0.211	0.090
	(0.86)	(1.15)	(0.68)	(1.08)	(0.64)
AbsDA	-0.004	-0.005	-0.003	0.963***	1.013***
	(-0.78)	(-0.99)	(-0.52)	(11.78)	(16.10)
G_growth	0.021	0.031	0.023	-0.196	0.181
	(0.42)	(0.53)	(0.46)	(-0.37)	(0.53)
Tax_growth	0.001	0.009	0.001	0.148	-0.157*
	(0.11)	(0.76)	(0.07)	(1.21)	(-1.85)
Constant	-0.106***	-0.074	-0.068**	-6.592***	-4.485***
	(-2.85)	(-1.52)	(-2.17)	(-18.10)	(-16.11)
Firm	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Year×Ind	Yes	No	No	No	No
N	13821	13821	13821	13821	13821
Adj.R2	0.04	0.02	0.01	0.15	0.29

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by company-level clustering (Cluster).

4.4. Big data tax enforcement and RPTs: Mechanism analysis

1. Mechanism 1: Reducing the information asymmetry between tax authorities and enterprises

According to our theoretical development, big data tax enforcement can help tax authorities to fully grasp the tax-related information of enterprises, alleviating the information asymmetry between tax authorities and enterprises and thus inhibiting major shareholders' tunneling activities by RPTs. Therefore, we expect that in enterprises with higher information asymmetry, big data tax enforcement is more likely to inhibit major shareholders' tunneling behaviors and reduce RPTs more.

As a crucial information intermediary in the capital market, analysts follow public companies and use their professional knowledge to interpret information related to the company's operations and release surplus forecast reports or valuation reports to the market, thus guiding investors' investment decisions. Analysts play a vital role in information collection and dissemination in the market, and analysts' forecasts often represent the market's assessment of the operating conditions of companies. Analysts'

forecasts are mainly based on their public and private information. The greater the forecast error and forecast disagreement among analysts, the less the outside world knows about the enterprise's operating conditions and the more serious the information asymmetry faced by firms (Lang et al., 2012; Balakrishnan et al., 2019).

To verify our expectations, we use analyst forecast error (*Error*) and analyst forecast disagreement (*Dispersion*) to measure the degree of firms' information asymmetry, respectively. We add the analyst forecast error *High_Error* and its interaction term with the *GTP* in model (1), and the regression results are shown in column (1) of Table 8. Similarly, we add the analyst forecast disagreement *High_Dispersion* and its interaction term with the *GTP* in model (1), and the regression results are shown in column (2) of Table 8. When the analyst forecast error is greater than the annual median, *High_Error* takes the value of 1, which indicates that the enterprise information asymmetry is higher, otherwise 0. When the analyst forecast disagreement is greater than the annual median, *High_Dispersion* takes the value of 1, which indicates that the enterprise information asymmetry is higher, otherwise 0. From Table 8 columns (1)-(2), we can find that the interaction terms of *GTP×High_Error* and *GTP×High_Dispersion* are all significantly negative at the 5% statistical level, indicating that the CTAIS-3 plays a greater role in reducing enterprise RPTs when the degree of enterprise information asymmetry is higher. The regression results in Table 8 imply that big data tax enforcement can alleviate the information asymmetry between tax authorities and enterprises, which is conducive to tax authorities to combat enterprises' profit-shifting behaviors, thus making enterprises reduce RPTs.

Table 8 Mechanism 1: Reducing information asymmetry between tax authorities and enterprises

	(1)	(2)
	RPT	RPT
GTP	-0.003	-0.002
	(-1.29)	(-0.63)
High_Error	0.001	
	(1.34)	
GTP×High_Error	**-0.007****	
	(-2.49)	

High_Dispersion		-0.001
		(-0.49)
GTP×High_Dispersion		**-0.008*** **
		(-2.71)
Size	0.004***	0.004***
	(2.94)	(2.66)
Lev	0.009*	0.007
	(1.80)	(1.28)
Growth	0.004***	0.004***
	(3.30)	(3.14)
Fcf	0.004	0.000
	(1.11)	(0.10)
Top1	0.030***	0.034***
	(3.46)	(3.54)
Top1_25	0.001***	0.001***
	(6.36)	(6.75)
Boardsize	0.017***	0.013**
	(3.03)	(2.14)
Indep	0.029*	0.020
	(1.96)	(1.27)
Ins_ratio	0.000	0.000
	(0.66)	(0.81)
Taxrate	0.020	0.015
	(1.10)	(0.77)
AbsDA	-0.003	-0.004
	(-0.54)	(-0.74)
G_growth	0.013	-0.008
	(0.25)	(-0.13)
Tax_growth	0.004	0.011
	(0.36)	(0.85)
Constant	-0.117***	-0.100***
	(-3.63)	(-2.87)
Firm	Yes	Yes
Year	Yes	Yes
N	13682	12222
Adj.R2	0.02	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by firm-level clustering (Cluster).

2. Mechanism 2: Improving the tax enforcement capability

Following theoretical development in this study, the second mechanism of the big

data tax enforcement's effect on the enterprises' RPTs is that the use of big data improves the tax data analysis ability of tax staffs, which in turn improves the tax enforcement capability. It is the most important feature that distinguishes big data tax enforcement from traditional tax enforcement.

The average education years of tax staff reflect the working ability of tax staffs and have an important influence on the tax enforcement capability. In addition, tax authorities mainly examine the reasonableness of transaction pricing of RPTs to determine whether there are tax evasions by enterprises, which involves some financial knowledge. Therefore, we use the financial knowledge of tax staffs to measure tax enforcement capability. We predict that in regions where tax staffs are less educated and have less financial knowledge, the data analysis capability possessed by the CTAIS-3 can compensate for the lack of tax staffs' working ability and improve tax enforcement ability of tax staffs, thus reducing RPTs. In order to verify the above prediction, we use the percentage of tax staffs with finance-related qualifications (i.e., CPA certificate and certified tax agents' certificate) divided by all tax staffs in the region as *CPAratio*, and the percentage of tax staffs at graduate level and above divided by all tax staffs as *Masterratio* to measure the tax enforcement ability, respectively. We add *Low_Masterratio* and its interaction term with *GTP* in model (1), and the regression results are shown in column (1) of Table 9; add *Low_CPAratio* and its interaction term with GTP in model (1), and the regression results are shown in column (2) of Table 9. When *Masterratio* is less than the annual median, *Low_Masterratio* takes the value of 1, which indicates that the regional tax staffs are less educated, otherwise 0. When *CPAratio* is less than the annual median, *Low_CPAratio* takes the value of 1, which indicates that the regional tax staffs are less financially literate, otherwise 0. From Table 9 column (1)-(2), it can be found that the regression coefficients of the interaction terms *GTP×Low_Masterratio* and *GTP×Low_CPAratio* are all significantly negative at the 10% statistical level, indicating that the CTAIS-3 plays a greater role in reducing the RPTs of enterprises when the tax enforcement capability of regional tax staffs is lower. The regression results in Table 8 indicate that big data

tax enforcement can make up for the lack of working ability of tax authorities, that is, modern information technology can improve the tax enforcement efficiency of tax staffs, thus enabling enterprises to reduce RPTs.

Table 9 Mechanism 2: Improving tax enforcement capacity of tax authorities

	Low education	Less CPA
	(1)	(2)
	RPT	RPT
GTP	-0.003	-0.004
	(-1.10)	(-1.62)
Low_Masterratio	-0.000	
	(-0.04)	
GTP×Low_Masterratio	**-0.007****	
	(-2.27)	
Low_CPAratio		-0.001
		(-0.72)
GTP×Low_CPAratio		**-0.005***
		(-1.95)
Size	0.003**	0.003***
	(2.46)	(2.61)
Lev	0.007	0.007
	(1.40)	(1.42)
Growth	0.004***	0.004***
	(3.50)	(3.43)
Fcf	0.004	0.004
	(1.01)	(1.01)
Top1	0.025***	0.025***
	(2.91)	(2.95)
Top1_25	0.001***	0.001***
	(6.63)	(6.63)
Boardsize	0.017***	0.017***
	(3.12)	(3.10)
Indep	0.026*	0.026*
	(1.79)	(1.77)
Ins_ratio	0.000	0.000
	(0.46)	(0.44)
Taxrate	0.026	0.026
	(1.43)	(1.41)
AbsDA	-0.003	-0.003
	(-0.57)	(-0.52)
G_growth	0.003	-0.004
	(0.06)	(-0.07)
Tax_growth	0.005	0.009

	(0.48)	(0.81)
Constant	-0.100***	-0.103***
	(-3.12)	(-3.22)
Firm	Yes	Yes
Year	Yes	Yes
N	13821	13821
Adj.R2	0.02	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by company-level clustering (Cluster).

4.5. Big data tax enforcement and RPTs: heterogeneity analysis

1. Big data tax enforcement and RPTs: from the perspective of motive and ability of large shareholders to tunneling

The prominent feature of the shareholding structure of China's listed companies is "one share is too big", i.e., the shareholding ratio of major shareholders is high. Large shareholders control a large number of subsidiaries through the pyramid structure, which leads to the separation of cash flow rights and control rights. The separation of the two rights allows the majority shareholder to achieve larger private gains at a low cost. Therefore, the higher the separation degree of the two rights, the greater the incentive for large shareholders to use their control rights to infringe on the interests of small and medium shareholders for maximizing their private interests. RPTs is an important way for large shareholders to transfer resources of listed companies and encroach on the interests of small and medium shareholders (Cheung et al., 2006). Hence, the higher the separation degree of two rights, the more motivated the major shareholders are to implement RPTs for tunneling the listed companies (Kang et al., 2014).

After the implementation of the CTAIS-3, the tax-related information supervision ability of tax authorities is stronger, that is, the tunneling behavior of large shareholders to transfer profits through RPTs is more easily to be detected and stopped by tax authorities. Given that the major shareholders with higher separation of two rights have stronger motivation to implement RPTs for tunneling, we expect that the inhibitory effect of the CTAIS-3 on RPTs of enterprises with stronger motives

for tunneling is more significant. In order to verify the above expectation, we add the large shareholder's motive of tunneling (*High_Motive*) and its interaction term with *GTP* in model (1). The regression results are shown in column (1) of Table 9. *High_Motive* takes the value of 1 when the separation of two rights is greater than the annual median, indicating that the major shareholder has a strong motive to tunneling, otherwise 0. The separation of two rights equals to the proportion of the listed company's control rights owned by the effective controller minus the proportion of the listed company's ownership rights. From column (1) of Table 9, we find that the coefficient of the interaction term *GTP×High_Motive* is significantly negative, indicating that the inhibitory effect of big data tax enforcement on RPTs is more prominent in enterprises with stronger motive of tunneling by major shareholders.

Moreover, we verify the baseline association from the perspective of tunneling ability. Due to their high shareholding, the major shareholders have the right to appoint executives to the listed companies (Claessen et al., 2000), which enhances their control over the business decisions of the companies and makes them capable of tunneling the listed companies through RPTs (Denis and McConnell, 2003). Therefore, we expect that after the implementation of the CTAIS-3, the decline of RPTs of enterprises with stronger tunneling ability of major shareholders is greater. Existing studies suggest that major shareholders other than the largest one can play a governance role in inhibiting controlling shareholders' self-interest behavior through RPTs and capital appropriation (Maury and Pajuste, 2005; Attig et al., 2009). Consequently, the ability of major shareholders to tunneling depends on their relative rights in the firm. We use the balance of shareholders equity structure (*Top1_25*) to measure the tunneling ability of large shareholders. Specifically, we add the large shareholder's tunneling ability (*High_Capacity*) and its interaction term with the *GTP* in model (1). The regression results are shown in column (2) of Table 9. *High_Capacity* takes a value of 1 when the equity checks and balances of the enterprise (Top1_25) takes a value greater than the annual median, indicating a stronger tunneling capacity of the major shareholders, otherwise 0. From column (2) of Table 9, we find that the

coefficient of the interaction term *GTP×High_Capacity* is significantly negative, implying that in enterprises with stronger tunneling capacity of the major shareholder, the negative correlation relationship between *GTP* and *RPT* is greater, i.e., the inhibitory effect of big data tax enforcement on RPTs is stronger. The regression results in Table 9 show that the big data tax enforcement has a greater inhibitory effect on the RPTs of enterprises with stronger tunneling motivation and tunneling ability, thus confirming the existence of RPTs' tunneling motive.

Table 9 Big data tax enforcement and RPTs: based on the perspective of large shareholders' tunneling

	Tunneling motivation	Tunneling ability
	(1)	(2)
	RPT	RPT
GTP	-0.002	-0.002
	(-0.92)	(-0.96)
High_Motive	0.009***	
	(4.76)	
GTP×High_Motive	**-0.009*** **	
	(-2.96)	
High_Capacity		0.004**
		(2.19)
GTP×High_Capacity		**-0.010*** **
		(-3.58)
Size	0.004***	0.003**
	(2.76)	(2.35)
Lev	0.008	0.006
	(1.55)	(1.25)
Growth	0.004***	0.004***
	(3.35)	(3.56)
Fcf	0.005	0.004
	(1.15)	(1.05)
Top1	0.023***	0.022**
	(2.58)	(2.39)
Top1_25	0.001***	0.001***
	(6.37)	(6.04)
Boardsize	0.017***	0.018***
	(2.94)	(3.15)
Indep	0.024	0.026*
	(1.57)	(1.77)
Ins_ratio	0.000	0.000
	(0.08)	(0.48)

Taxrate	0.019	0.027
	(1.00)	(1.47)
AbsDA	-0.003	-0.002
	(-0.62)	(-0.45)
G_growth	0.017	0.001
	(0.32)	(0.02)
Tax_growth	0.008	0.006
	(0.72)	(0.55)
Constant	-0.111***	-0.097***
	(-3.38)	(-3.04)
Firm	Yes	Yes
Year	Yes	Yes
N	13351	13821
Adj.R2	0.02	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by firm-level clustering (Cluster).

2. Big data tax enforcement and RPTs: based on the perspective of corporate governance

Although large shareholders have the motivation and ability to seek for private gain at the expense of other shareholders (Cheung et al., 2009; Lei and Song, 2011), fortunately, good corporate governance mechanisms can act as a watchdog over the behavior of large shareholders and thus discourage them from tunneling (Stein and Zhao, 2019). Usman et al. (2022) found that female directors reduced the scale of RPTs and led to higher operating performance. Chen et al. (2012) found that fierce product market competition makes companies facing large shareholder tunneling more vulnerable to operating difficulties and therefore helps to curb large shareholder tunneling and plays a role of external governance. Correspondingly, we expect with a poor corporate governance of enterprises, the tunneling by major shareholders is serious. By this prediction, the big data tax enforcement plays a greater role and the relationship between the CTAIS-3 and RPTs will be more significant.

Corporate governance is measured from both internal and external perspectives in this study. Drawing on Bai et al. (2005) and Li et al. (2018), we construct a corporate

governance index (*CG_Index*) to measure the level of internal corporate governance[8] and adopts the product market competition faced by firms to measure external corporate governance. To verify the above expectation, we add the internal governance *Low_CG* and its interaction term with the *GTP* in model (1), and the regression results are shown in column (1) of Table 10. Similarly, we add the external governance *Low_Competition* and its interaction term with the *GTP* in model (1), and the regression results are shown in column (2) of Table 10. When the corporate governance index (*CG_Index*) of the enterprise is less than the annual median, *Low_CG* takes the value of 1, which indicates poor corporate governance, and 0 otherwise; when the Herfindahl index of the industry in which the enterprise is located is higher than the annual median, *Low_Competition* takes the value of 1, which indicates a low degree of product market competition, and 0 otherwise. By Table 10 columns (1)-(2) it can be found that the regression coefficients of the interaction terms *GTP×Low_CG* and *GTP×Low_Competition* are both significantly negative at the 5% statistical level, indicating that the negative correlation between *GTP* and *RPT* is more significant when the corporate governance of enterprises are poor, i.e., the inhibitory effect of big data tax enforcement on RPTs is stronger. The regression results in Table 10 indicate that the inhibitory effect of big data tax enforcement on RPTs is influenced by the corporate governance status of enterprises themselves, and there is a substitution relationship between good corporate governance and big data tax enforcement.

Table 10 Big data tax enforcement and RPTs: a corporate governance-based perspective

	Internal Governance	External Governance
	(1)	(2)
	RPT	RPT
GTP	-0.002	-0.002
	(-0.68)	(-1.03)

[8] We use eight variables for principal component analysis, including the shareholding ratio of the first largest shareholder, the square sum of the shareholding ratios of the second to tenth largest shareholders, the size of the directors, the proportion of independent directors, the duality of Chairman and CEO, cross-listing, the nature of ownership, and the proportion of management shareholding. The first principal component was chosen as the corporate governance index, with a higher value indicating a higher level of corporate governance.

Low_CG	0.006***	
	(3.39)	
GTP×Low_CG	**-0.007****	
	(-2.54)	
Low_Competition		-0.000
		(-0.13)
GTP×Low_Competition		**-0.007****
		(-2.37)
Size	0.003**	0.003***
	(2.56)	(2.62)
Lev	0.007	0.007
	(1.40)	(1.46)
Growth	0.003***	0.004***
	(2.72)	(3.41)
Fcf	0.004	0.004
	(1.02)	(1.09)
Top1	0.020**	0.026***
	(2.26)	(2.96)
Top1_25	0.001***	0.001***
	(5.83)	(6.60)
Boardsize	0.014**	0.017***
	(2.48)	(3.10)
Indep	0.029*	0.027*
	(1.90)	(1.81)
Ins_ratio	0.000	0.000
	(0.88)	(0.45)
Taxrate	0.011	0.026
	(0.62)	(1.44)
AbsDA	-0.005	-0.003
	(-0.88)	(-0.51)
G_growth	-0.051	0.005
	(-0.98)	(0.10)
Tax_growth	0.013	0.008
	(1.19)	(0.68)
Constant	-0.093***	-0.105***
	(-2.88)	(-3.28)
Firm	Yes	Yes
Year	Yes	Yes
N	13194	13821
Adj.R2	0.02	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are

adjusted by firm-level clustering (Cluster).

3. Big data tax enforcement and RPTs: based on the perspective of tax avoidance

Although tax avoidance activities bring tax savings for a company, tax planning activity can also be used as an excuse for insiders to tunnel profits out from the listed company to its parent company or other related companies (Wong et al., 2015). Besides, aggressive tax avoidance can also facilitate the diversion of corporate resources by firm insiders through the additional cash flows from tax savings and the increased financial opacity from tax planning. Thus, firms with higher levels of tax avoidance tend to have more insider benefit appropriation. We predict that the relationship between the CTAIS-3 and RPTs is more significant in firms with a higher degree of tax avoidance.

In order to verify the above expectation, we add *High_ETR* and its interaction term with *GTP* in model (1). The regression results are shown in Table 11. When the average tax avoidance degree of enterprise before the implementation of CTAIS-3 is greater than the median of the sample[®], *High_ETR* takes the value of 1, indicating that the enterprise has a higher tax avoidance degree, otherwise 0. It can be found that the regression coefficient of the interaction term *GTP*×*High_ETR* is significantly negative at the 1% statistical level, indicating that the negative relationship between *GTP* and *RPT* is more significant in enterprise with a higher degree of tax avoidance before the implementation of the CTAIS-3, i.e., the inhibitory effect of big data tax enforcement on RPTs is stronger.

Table 11 Big data tax enforcement and RPTs: based on the perspective of tax avoidance

	(1)
	RPT
GTP	-0.002
	(-0.88)
High_ETR	0.000
	(.)

[®] Referring to the practice of Desai & Dharmapala (2006), we use the tax difference (i.e., the difference between pre-tax accounting income and taxable income) to return to the total accrual profit to obtain the regression residual to measure the degree of corporate tax avoidance ETR. This regression residual removes the part of the tax variance explained by earnings management and is a more accurate measure of corporate tax evasion.

GTP×High_ETR	**-0.009*** **
	(-3.06)
Size	0.003***
	(2.65)
Lev	0.006
	(1.18)
Growth	0.004***
	(3.59)
Fcf	0.005
	(1.16)
Top1	0.026***
	(3.02)
Top1_25	0.001***
	(6.52)
Boardsize	0.018***
	(3.15)
Indep	0.027*
	(1.83)
Ins_ratio	0.000
	(0.28)
Taxrate	0.023
	(1.20)
AbsDA	-0.003
	(-0.49)
G_growth	0.003
	(0.06)
Tax_growth	0.008
	(0.69)
Constant	-0.108***
	(-3.33)
Firm	Yes
Year	Yes
N	13627
Adj.R2	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by firm-level clustering (Cluster).

4. Big Data Tax enforcement and RPTs: based on the perspective of tax enforcement motivation and information technology skill

Previous studies have shown that the imbalance between fiscal revenues and fiscal expenditures is a major problem for local governments in transition economies

(Bird and Vaillancourt, 1999). According to Chen et al. (2011), during the period 1993-2008, the fiscal expenditure of local governments in China was 1.375 times higher than the fiscal revenue. Local governments have strong incentives to increase tax revenues to cover fiscal deficits. For example, Chen (2016) found that the abolition of agricultural tax brought fiscal pressure on county governments, which prompted them to strengthen tax enforcement to combat corporate tax evasion. Therefore, we expect that in regions with higher fiscal pressure (higher fiscal deficits), local governments have stronger incentives to use the CTAIS-3 to increase tax revenue to cover fiscal deficits, that is, the disincentive effect of big data tax enforcement on RPTs is greater. Referring to existing studies (Tang et al., 2017), we adopt the ratio of the difference between provincial government's fiscal expenditure and fiscal revenue in the current year to GDP in the current year to measure local government fiscal pressure (*Deficit*). The larger the value is, the greater the government fiscal pressure and the stronger the incentive for local governments to strengthen tax enforcement are.

To verify the above expectation, we add *High_Deficit* and its interaction term with *GTP* in model (1). The regression results are shown in column (1) of Table 11. *High_Deficit* takes the value of 1 when the local government fiscal pressure (*Deficit*) is greater than the annual median, indicating that the local government fiscal pressure is higher, otherwise 0. It can be found that the regression coefficient of the interaction term *GTP*×*High_Deficit* is significantly negative at the 5% statistical level, indicating that when the local government fiscal pressure is higher, the negative correlation between *GTP* and *RPT* is more significant, i.e., the inhibitory effect of big data tax enforcement on RPTs is stronger.

As the CTAIS-3 system integrates modern information technologies such as big data and cloud computing, the information technology skill of tax staffs has an important influence on whether the role of the CTAIS-3 can be effectively played. We expect that the CTAIS-3 functions more effectively when the information technology skill of tax staffs is higher, thus exerting a greater inhibiting effect on RPTs. To verify the above expectation, we add *High_Computer* and its interaction term with *GTP* in

model (1), and the regression results are shown in column (2) of Table 11. *High_Computer* takes the value of 1 when the percentage of the number of tax staffs with computer skill level 2 or above certificates is greater than the annual median, indicating that the local tax staffs has a high IT skill level, otherwise 0. It can be found that the regression coefficient of the interaction term *GTP×High_Computer* is significantly negative at the 10% statistical level, indicating that the negative correlation between *GTP* and *RPT* is more significant in areas where the information technology skill of tax staffs is higher.

The regression results in Table 11 indicate that the inhibitory effect of big data tax enforcement on RPTs depends on the local government's motivation to collect taxes and the ability of local tax staffs to effectively use the CTAIS-3.

Table 11 Big Data Tax enforcement and RPTs: Based on the Perspective of Tax Motivation and Information Technology Skill

	Tax motivation	Ability to use the CTAIS-3
	(1)	(2)
	RPT	RPT
GTP	0.002	-0.005**
	(1.02)	(-2.16)
High_Deficit	-0.001	
	(-0.17)	
GTP×High_Deficit	**-0.008****	
	(-1.96)	
High_Computer		-0.001
		(-0.72)
GTP×High_Computer		**-0.005***
		(-1.70)
Size	0.003	0.003***
	(1.31)	(2.68)
Lev	0.005	0.007
	(0.74)	(1.40)
Growth	0.004***	0.004***
	(2.73)	(3.60)
Fcf	0.004	0.003
	(1.04)	(0.77)
Top1	0.026	0.026***
	(1.47)	(2.93)

Top1_25	0.001***	0.001***
	(3.17)	(6.76)
Boardsize	0.017*	0.016***
	(1.81)	(2.87)
Indep	0.027	0.022
	(1.16)	(1.51)
Ins_ratio	0.000	0.000
	(0.44)	(0.36)
Taxrate	0.024	0.026
	(0.81)	(1.42)
AbsDA	-0.003	-0.004
	(-0.49)	(-0.69)
G_growth	-0.003	-0.001
	(-0.04)	(-0.03)
Tax_growth	0.005	0.004
	(0.36)	(0.34)
Constant	-0.098	-0.101***
	(-1.64)	(-3.16)
Firm	Yes	Yes
Year	Yes	Yes
N	13821	13667
Adj.R2	0.02	0.02

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by firm-level clustering (Cluster).

4.6. Motives of corporate RPTs: tax avoidance or tunneling?

The previous conclusion shows that the scale of corporate RPTs decreases significantly after the implementation of CTAIS-3. But how does it affect corporate value? According to existing studies, on the one hand, the major shareholder's tunneling is the main motive for RPTs (Cheung et al., 2006). Therefore, RPTs will damage the value of the company because of underlying tunneling activities. On the other hand, if there is a significant difference in the corporate income tax rate applied to different enterprises in the same group, enterprises can wholly reduce the tax burden of the group through RPTs, thus increasing the value of the company (Lo et al., 2006). Nevertheless, is the motive of tunneling or tax avoidance dominant in the RPTs of Chinese listed companies? In order to answer this question, we further examine

whether the decline of RPTs brought by big data tax enforcement facilitate to increase firm value.

We examine both operating performance (*Gross_sales*) and market performance (*Aj_Ret*), presenting the regression results in Table 12. Operating performance is measured by gross operating margin[®], which equals to the difference between operating revenues and operating costs divided by operating revenues. Market performance is measured by market-adjusted stock returns, which equals to the difference of company's individual stock returns minus the market returns. *High_PreRPT* is a dummy variable for RPTs, takes the value of 1 when the mean value of RPT in the years before the implementation of CTAIS-3 is greater than the median value of the sample, which means that the scale of RPTs of enterprises is larger, otherwise 0. From Table 12, we find that the regression coefficients of *High_PreRPT×GTP* are all significantly positive at the 1% statistical level, which indicates that in the enterprises with larger scale of RPTs before the implementation of CTAIS-3, the implementation of CTAIS-3 is more beneficial to improve their operating performance and market performance. As enterprises with larger scale of RPTs before the implementation of the CTAIS-3 have a greater decrease in the scale of RPTs, the results in Table 12 indicate that big data tax enforcement reduces the RPTs behaviors of listed companies with opportunistic tendencies, which is conducive to improving enterprise value. The results provide evidence that the motive of tunneling by RPTs of Chinese listed companies plays a dominant role.

Table 12 Motives of corporate RPTs: tax avoidance or tunneling?

	(1) Gross_sales	(2) Aj_Ret
GTP	-0.001	0.029
	(-0.20)	(1.27)
High_PreRPT×GTP	**0.011*** **	**0.097*** **

® The reason for using gross operating margin to measure operating performance is that: the main way for major shareholders to tunneling through RPTs of goods and services is to buy from listed companies at low prices or sell at high prices, if the CTAIS-3 improves operating performance by suppressing RPTs, the most direct manifestation is the increase of enterprise gross operating margin; net profit is more manipulated, while gross operating margin is more difficult to manipulate. For the sake of robustness, this paper also uses total net asset margin to measure operating performance, and the conclusion still holds.

	(2.69)	(3.38)
High_PreRPT	0.000	0.000
	(.)	(.)
Size	0.044***	0.467***
	(24.57)	(37.90)
Lev	-0.136***	0.192***
	(-19.89)	(4.11)
Growth	-0.006***	0.032***
	(-4.04)	(3.12)
Fcf	0.026***	0.159***
	(4.62)	(3.84)
Top1	0.068***	-0.161*
	(5.47)	(-1.89)
Top1_25	-0.000	0.002**
	(-1.19)	(2.11)
Boardsize	0.008	-0.315***
	(0.97)	(-5.73)
Indep	-0.009	-0.477***
	(-0.43)	(-3.28)
Ins_ratio	0.000	0.002
	(0.74)	(1.61)
Taxrate	0.017	0.249
	(0.65)	(1.38)
AbsDA	0.038***	0.020
	(4.81)	(0.37)
G_growth	0.139*	-1.442***
	(1.91)	(-2.89)
Tax_growth	-0.013	0.620***
	(-0.79)	(5.56)
Constant	-0.724***	-10.040***
	(-15.69)	(-31.73)
Firm	Yes	Yes
Year	Yes	Yes
N	13759	13531
Adj.R2	0.10	0.18

Note: (1) Values reported in parentheses are t-statistics; (2) "*", "**" and "***" indicate 10%, 5% and 1% significance levels, respectively. (3) Standard errors are adjusted by firm-level clustering (Cluster).

5. Conclusion

The rapid development of big data has profoundly changed people's life, work and way of thinking. In order to improve the efficiency of tax enforcement, the State

Taxation Administration of China employed the "Internet + Taxation" strategy, implementing the CTAIS-3 on a pilot basis in 2013 and extending it nationwide by the end of 2016. The CTAIS-3 integrates information technologies such as big data, cloud computing, and the Internet to greatly improve the ability of tax authorities to supervise tax-related information, effectively combat tax evasion by enterprises (Xiao and Shao, 2020), and improve surplus information content (Zhao, 2021) and investment efficiency (Zhang et al., 2022). However, it is unclear whether the use of big data can facilitate the corporate governance role of tax enforcement. Given firms often engage in a large number of RPTs for the motive of tunneling (Cheung et al., 2006; Lo et al., 2010), this paper examines the ways in which big data tax enforcement plays a governance role from the perspective of RPTs, a specific business behavior of enterprises.

Basing on the theoretical analysis of the impact of big data tax enforcement on RPTs, this study conducts an empirical test taking the implementation of the CTAIS-3 as a quasi-natural experiment, uses A-share listed companies from 2008 to 2016 as the research sample, and finds that the CTAIS-3 significantly inhibits the RPTs of enterprises. The analysis of the mechanism indicates that the CTAIS-3 works mainly by reducing the information asymmetry between tax authorities and enterprises and improving the tax enforcement ability of tax authorities. The heterogeneity analysis finds that the big data tax enforcement plays a greater role on RPTs in enterprises with stronger controlling shareholders' motive and ability to tunneling, enterprises with worse corporate governance, enterprises with stronger motive to tax avoidance, and enterprises in regions with stronger tax enforcement motive of tax staffs and better information technology skill. Further analysis finds that enterprises with a higher proportion of RPTs before the implementation of the CTAIS-3 better increase their enterprise value with the implementation of the CTAIS-3, confirming that the motive of tunneling plays a dominant role in the RPTs of Chinese listed companies.

This study has strong theoretical and practical values. First, from the perspective of the specific business behavior of enterprises, RPTs, this study empirically examines

the ways in which big data tax enforcement affects the behavior of enterprises. It finds that the use of big data in tax enforcement can significantly improve the efficiency of tax enforcement and inhibit the RPTs behavior of enterprises with tunneling motive and tax avoidance motive, thus playing a role of safeguarding government tax revenue and improving corporate governance. It highlights the importance of big data in government regulation and provides research reference for subsequent big data regulation. Second, this study finds that big data can empower tax enforcement mainly by reducing information asymmetry and improving decision analysis, indicating that the rapidly development of modern information technology greatly enhances the accessibility of information, improves the efficiency of economic agents' decision making, and boosts the quality of economic operation. Government departments should accelerate the pace of digital transformation and make full use of modern information technology such as big data to improve the efficiency of government supervision. In addition, this study finds that the role of the CTAIS-3 depends on the information technology skill of tax staffs. Therefore, government departments should gradually adjust the staff structure and increase the training of information technology knowledge for existing staffs to lay the foundation for the digital transformation of government supervision. Third, we find that cracking down on RPTs is conducive to improving corporate value, suggesting that the motive of tunneling by RPTs plays a dominate role among Chinese listed companies. This provides a policy reference for regulators to take actions against RPTs fraud of enterprises on the one hand and warns investors to be alert to RPTs of listed companies to make scientific investment decisions on the other hand.

Reference

Almeida, H. V., & Wolfenzon, D. (2006). A theory of pyramidal ownership and family business groups. The journal of finance, 61(6), 2637-2680.

Anderson, M. C., Banker, R. D., & Ravindran, S. (2006). Value Implications of Investments in Information Technology. Management Science, 52(9), 1359-1376.

Attig, N., El Ghoul, S., & Guedhami, O. (2009). Do multiple large shareholders play a corporate governance role? Evidence from East Asia. Journal of Financial Research, 32(4), 395-422.

Balakrishnan, K., Blouin, J. L., & Guay, W. R. (2019). Tax aggressiveness and corporate transparency. The Accounting Review, 94(1), 45-69.

Bauer, A. M., Fang, X., & Pittman, J. A. (2021). The importance of IRS enforcement to stock price crash risk: The role of CEO power and incentives. The Accounting Review, 96(4), 81-109.

Beck, T., Levine, R., & Levkov, A. (2010). Big bad banks? The winners and losers from bank deregulation in the United States. The Journal of Finance, 65(5), 1637-1667.

Bennouri, M., Nekhili, M., & Touron, P. (2011). Does Auditors' Reputation "Discourage" Related Party Transactions. The French Case, 46-68.

Berkman, H., Cole, R. A., & Fu, L. J. (2010). Political connections and minority-shareholder protection: Evidence from securities-market regulation in China. Journal of financial and quantitative analysis, 45(6), 1391-1417.

Bertrand, M., & Mullainathan, S. (2003). Enjoying the quiet life? Corporate governance and managerial preferences. Journal of political Economy, 111(5), 1043-1075.

Besley, T., & Persson, T. (2012). Public finance and development. Draft Chapter for the Handbook of Public Economics.

Bird, R. M. and F. Vaillancourt. 1999. Fiscal decentralization in developing countries. New York, NY: Cambridge University Press.

Brazel, J. F., & Dang, L. (2008). The effect of ERP system implementations on the management of earnings and earnings release dates. Journal of information systems, 22(2), 1-21.

Cai, H., & Liu, Q. (2009). Competition and corporate tax avoidance: Evidence from Chinese industrial firms. The Economic Journal, 119(537), 764-795.

Chen, F., Hope, O.-K., Li, Q., & Wang, X. (2011). Financial reporting quality and investment efficiency of private firms in emerging markets. The accounting review, 86(4), 1255-1288.

Chen, S., Wang, K., & Li, X. (2012). Product market competition, ultimate controlling structure and related party transactions. China Journal of Accounting Research, 5(4), 293-306.

Cheung, Y.-L., Rau, P. R., & Stouraitis, A. (2006). Tunneling, propping, and expropriation: Evidence from connected party transactions in Hong Kong. Journal of Financial economics, 82(2), 343-386.

Cheung, Y. L., Jing, L., Tong, L., Rau, R., & Stouraitis, A. (2009). Tunneling and propping up: An analysis of related party transactions by Chinese listed companies. Social Science Electronic Publishing, 17(3), 372-393.

Claessens, S., Djankov, S., Fan, J., & Lang, L. (2000). Expropriation of minority shareholders in East Asia: Institute of Economic Research, Hitotsubashi University.

Cook, K. S. (1977). Exchange and power in networks of interorganizational relations. The sociological quarterly, 18(1), 62-82.

Dechow, P. M., & Dichev, I. D. (2002). The quality of accruals and earnings: The role of accrual estimation errors. The accounting review, 77(s-1), 35-59.

DeFond, M. L., Hung, M., Li, S., & Li, Y. (2015). Does mandatory IFRS adoption affect crash risk? The Accounting Review, 90(1), 265-299.

Denis, D. K., & McConnell, J. J. (2003). International corporate governance. Journal of financial and quantitative analysis, 38(1), 1-36.

Desai, M. A., & Dharmapala, D. (2006). Corporate tax avoidance and high-powered incentives. Journal of financial Economics, 79(1), 145-179.

Desai, M. A., Dyck, A., & Zingales, L. (2007). Theft and taxes. Journal of financial economics, 84(3), 591-623.

Dyck, A., & Zingales, L. (2004). Private benefits of control: An international comparison. The journal of finance, 59(2), 537-600.

Fang, J., Pittman, J., Zhang, Y., & Zhao, Y. (2017). Auditor choice and its

implications for group-affiliated firms. Contemporary Accounting Research, 34(1), 39-82.

Fauver, L., Hung, M., Li, X., & Taboada, A. G. (2017). Board reforms and firm value: Worldwide evidence. Journal of Financial Economics, 125(1), 120-142.

Fisman, R., & Khanna, T. (2004). Facilitating development: The role of business groups. World Development, 32(4), 609-628.

Guedhami, O., & Pittman, J. (2008). The importance of IRS monitoring to debt pricing in private firms. Journal of Financial Economics, 90(1), 38-58.

Hanlon, M., Hoopes, J. L., & Shroff, N. (2014). The effect of tax authority monitoring and enforcement on financial reporting quality. The Journal of the American Taxation Association, 36(2), 137-170.

Hitt, L. M., Wu, D. J., & Zhou, X. (2002). Investment in enterprise resource planning: Business impact and productivity measures. Journal of management information systems, 19(1), 71-98.

Hoopes, J. L., Mescall, D., & Pittman, J. A. (2012). Do IRS audits deter corporate tax avoidance? The accounting review, 87(5), 1603-1639.

Hope, O.-K., Lu, H., & Saiy, S. (2019). Director compensation and related party transactions. Review of Accounting Studies, 24(4), 1392-1426.

Jian, M., & Wong, T. J. (2010). Propping through related party transactions. Review of Accounting Studies, 15(1), 70-105.

Jiang, H., Tian, G., & Zhou, D. (2021). The influence of the deregulation of short-selling on related-party transactions: Evidence from China. Journal of Business Finance & Accounting, 48(5-6), 1022-1056.

Kang, M., Lee, H.-Y., Lee, M.-G., & Park, J. C. (2014). The association between related-party transactions and control–ownership wedge: Evidence from Korea. Pacific-Basin Finance Journal, 29, 272-296.

La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (1999). Corporate ownership around the world. The journal of finance, 54(2), 471-517.

Lang, M., Lins, K. V., & Maffett, M. (2012). Transparency, liquidity, and valuation:

International evidence on when transparency matters most. Journal of Accounting Research, 50(3), 729-774.

Lei, A. C., & Song, F. M. (2011). Connected transactions and firm value: Evidence from China-affiliated companies. Pacific-Basin Finance Journal, 19(5), 470-490.

Lennox, C. S., & Wu, X. (2022). Mandatory internal control audits, audit adjustments, and financial reporting quality: Evidence from China. The Accounting Review, 97(1), 341-364.

Li, J., Wang, X., & Wu, Y. (2020). Can government improve tax compliance by adopting advanced information technology? Evidence from the Golden Tax Project III in China. Economic modelling, 93, 384-397.

Liu, Q., & Lu, Y. (2015). Firm Investment and Exporting: Evidence from China's Value -Added Tax Reform. Journal of International Economics, 97(2), 392-403.

Lo, A. W., Wong, R. M., & Firth, M. (2010). Tax, financial reporting, and tunneling incentives for income shifting: An empirical analysis of the transfer pricing behavior of Chinese-listed companies. Journal of the American Taxation Association, 32(2), 1-26.

Masli, A., Peters, G. F., Richardson, V. J., & Sanchez, J. M. (2010). Examining the potential benefits of internal control monitoring technology. The Accounting Review, 85(3), 1001-1034.

Maury, B., & Pajuste, A. (2005). Multiple large shareholders and firm value. Journal of Banking & Finance, 29(7), 1813-1834.

Pincus, M., Tian, F., Wellmeyer, P., & Xu, S. X. (2017). Do clients' enterprise systems affect audit quality and efficiency? Contemporary Accounting Research, 34(4), 1975-2021.

Shleifer, A., & Vishny, R. W. (1997). A survey of corporate governance. The journal of finance, 52(2), 737-783.

Stein, L. C., & Zhao, H. (2019). Independent executive directors: How distraction affects their advisory and monitoring roles. Journal of Corporate Finance, 56, 199-223.

Tang, T., Mo, P. L. L., & Chan, K. H. (2017). Tax collector or tax avoider? An

investigation of intergovernmental agency conflicts. The Accounting Review, 92(2), 247-270.

Usman, M., Gull, A. A., Zalata, A. M., Wang, F., & Yin, J. (2022). Female board directorships and related party transactions. British Journal of Management, 33(2), 678-702.

Wong, R. M., Kim, J. B., & Lo, A. W. (2015). Are related-party sales value-adding or value-destroying? Evidence from China. Journal of International Financial Management & Accounting, 26(1), 1-38.

Xiao, C., & Shao, Y. (2020). Information system and corporate income tax enforcement: Evidence from China. Journal of Accounting and Public Policy, 39(6), 106772.

Zhang, L., Chen, W., & Peng, L. (2022). The impact of tax enforcement on corporate investment efficiency: evidence from the tax administration information system. Accounting & Finance.

Zhao, L. (2021). The Effect of Tax Authority Enforcement on Earnings Informativeness. European Accounting Review, 1-20.

Zhu, C. (2019). Big data as a governance mechanism. The Review of Financial Studies, 32(5), 2021-2061.

Bai, C., Liu, Q., Lu, Z., Song, M., & Zhang, J. (2005). An empirical study on the governance structure of Chinese listed companies Economic Research (02), 81-91.

Chen, X. (2016). Fiscal pressure, tax collection and management and regional inequality Chinese Social Sciences (04), 53-70 + 206.

Li, C., Liu, B., Zhou, P., & Zhang, X. (2018). The stone of its mountain: QFII and information disclosure of listed companies Financial Research (12), 138-156.

Zhang, K., Ou, Y., & Li, W. (2020). Why it is difficult to reduce taxes: information technology, tax ability and enterprise tax evasion Economic research, 55 (03), 116-132.