The London School of Economics and Political Science

Essays on the Dispersion of Effective VAT Rates in China: Causes and Consequences

Shawn Xiaoguang Chen

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, June 2015

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 67,052 words.

Acknowledgement

This work cannot be completed without enormous helps and supports. In the first place, I would express my genuine gratitude to Francesco Caselli and Camille Landais for their attentive supervision. They are always approachable, encouraging, enlightening. I am also grateful to Tim Besley for his insightful comments and suggestions that constantly make big differences.

I have also benefited substantially from discussions with a long list of professors and visitors at LSE, as well as those at various conferences, including Henrik Kleven, Johanness Spinnewijn, Robin Burgess, Miguel Almunia, Torsten Persson, Roger Gordon, Frank Cowell, Ethan Ilzetzki, Steve Pischke, Michael Peters, Rachel Ngai, John Sutton, Gabriel Zucman, Alan Manning, Alwyn Young, Oriana Bandiera, Guy Michaels, Ronny Razin, Keyu Jin, Mark Schankerman, Pasquale Schiraldi, Athar Hussain, Ray Fisman, Chang-Tai Hsieh, Pete Klenow, Michael Devereux, Jacob Svensson, Chenggang Xu, Peter Robertson. Considerable supports also come from professors in China, including Chiwa Yuen, Qing Zhu, Shumin Yue, Qingwang Guo, Liutang Gong, Ximing Yue, Heng Yin, Hengfu Zou, Xiaoyong Cui, Bingyang Lv, Junxue Jia, Jing Zhang, Xiaolin Long, Xixi Lin. For Financial Support, I thank the International development Research Council in Canada.

I am a lucky man to have friends and officemates around at LSE spraying colours along my journey on “dismal science”. They are Eddy Tam, Yu-Hsiang Lei, Yi Fan, Min Zhang, Shiyang Huang, Michael Best, Anne Brockmeyer, Amanda Fitzgerald, Mohammad Vesal, Mazhar Waseem, Albert Brue-Perez, Florian Blum, Anders Jensen, Xuezhu Shi, Panos Mavrokonstantis, Milad Khatib-Shahidi. I am also grateful to massive feedbacks and comments by participants at Work-in-Progress seminars at LSE on Macro, Labour, Public Economics, Development and Growth.

The last but the deepest acknowledgement is given to my family, particularly to my mum. From them I unfailingly feel love, learn tolerance, and borrow courage.

Abstract

It is well known that tax administration can be subject to an influence of political power, and bad tax administration may lead to an efficiency loss. However, both the extent and the mechanisms of the political intervention and the efficiency loss are still not fully understood in empirical works.

Using the Chinese Annual Survey of Manufacturing Firms, digitized data on the turnover of prefectural secretaries of the Chinese Communist Party, and the County Public Finance Statistics Yearbook in China from the year 2000 to 2007, the three chapters in this Ph.D. thesis aim to contribute to our understanding of following three questions: 1. How do local government incentives affect tax enforcement and effective tax rate of VAT? 2. What is the role of local politicians in selective tax enforcement across industries? 3. To what extent does the dispersion in the effective VAT rate across firms lead to production efficiency loss via the channel of resource misallocation?

The results suggest that: 1. Weak local government incentives, rather than lack of information on tax base, lead to a low effective VAT rate in China. 2. There is an increasing favouritism in tax enforcement towards capital-intensive industries as the prefectural secretaries of the Chinese Communist Party stay longer in office. On the contrary, labour-intensive industries face tougher tax enforcement. 3. A tax-neutral reform which eliminates the dispersion in VAT rates across firms in the same 4-digit industry produces a gain in aggregate TFP in the order of 7.9% of GDP on average in the period from 2000 to 2007.

Table of Contents

Chapter 1

The Effect of a Fiscal Squeeze on Tax Enforcement:

Evidence from a Natural Experiment in China*

Abstract

How significantly, and through what mechanisms, can taxation be constrained by political incentives? Which factors can shape the local political incentives in institutions with multi-level governments? This paper uses the abolition of the agricultural tax in 2005 across China as a natural 'fiscal squeeze' experiment to address these questions. I show that the revenue loss of county governments were largely and quickly offset by tougher tax enforcement. In particular, counties increased effective VAT rates without substantial changes in statutory VAT rates and changes in the reported tax base. This suggests that taxation capacity is not necessarily binding and constrained by access to information on tax base, as has been highlighted in many studies. There is a potentially large role for governments to endogeneously determine on how aggressively to enforce the statutory taxes. The incentive for VAT enforcement can be weakened, however, if the county: (1) receives a lower proportion of total VAT revenues after sharing with prefectural and provincial governments; (2) has a broader VAT tax base; or, (3) has more abundant sources of revenue from land sales. These findings are consistent with a model of endogeneous tax enforcement in presence of politicians' personal interests and tax competition in an institution of multi-level governments.

*The author is grateful to Francesco Caselli and Camille Landais for their supervision of this project and other related topics, and especially thankful to Tim Besley, Kleven Kleven, and John Sutton for their generous help and comments. Acknowledgements go to Johannes Spinnewijn, Torsten Persson, Ethan Ilzetzki, Steve Pischke, Robin Burgess, Frank Cowell, Ronny Razin, Guy Michaels, Ximin Yue for their suggestions. The author also appreciates the feedback from participants in Work-in-Progress Seminars on Public Economics, Macroeconomics and Labour at LSE, and benefits from discussion with Qing Zhu, Shumin Yue, Qingwang Guo, Xiaoyong Cui, Bingyang Lv, Junxue Jia, and Jing Zhang. Special thanks to Heng Yin and Liutang Gong for data supports. The author, of course, bears the responsibility for any errors in the paper.

1 Introduction

Poor countries are fiscally characterized by lower tax-to-GDP ratios, collecting less than 20 per cent of GDP, while rich countries take up to half of their GDP in tax revenue (Gordon and Li, 2009; Kleven, 2014; Besley and Torsten, 2014).

The reason why the tax-to-GDP ratio is low in poor countries continues to be in debate. Recent theoretical studies argue that it can simply be a result of a lack of information. Underlying causes may include informality, cash transactions, and the underground economy (Gordon and Li, 2009), or a lack of third-party reporting by large modern firms (Kleven, et al., 2009). The role of information is also emphasized by a growing empirical literature based on field experiments or quasi-experiments, which highlight the importance of tax administration and the role of third parties, suppliers, and consumers in revealing information (Slemrod et al., 2001; Kleven et al., 2011; Pomeranz, 2013; Naritomi, 2013).

Another strand of literature argues that real world tax administration is not only constrained by a society's economic structure, but also determined by incentives shaped by political institutions and power structures (Besley and Persson, 2013). Although the importance of government taxation incentives has long been widely recognized, we still lack a rigorous empirical understanding of which incentives are effective, and through what mechanisms they operate.[1]

This paper exploits the nation-wide abolition of the agricultural tax in China in 2005 as a natural 'fiscal squeeze' experiment in which county governments suffered a severe revenue loss. To meet their ongoing normal commitments to finance local services, county governments were forced to raise public revenue. Using the cross county variations in revenue loss has allows me to apply the standard diff-in-diff approach to identify the effect of local governments' incentives to enforce tax.

The results show that the agricultural tax revenue loss can be largely offset by tougher tax enforcement on other existing taxes in the same year of abolition. On average, the total tax revenue net of agricultural tax rose by 10%. And total tax revenue was virtually unchanged. This suggests that Chinese counties do not face binding fiscal capacity constraints. [2]

To understand the role of incentives, this research particularly focuses on formal manufacturing firms and their effective VAT rate, which is defined as the ratio of a firm's payable VAT to their reported valued-added. This is distinct from the statutory VAT

[1] Khwaja et al. (2014) provide a recent influential study in this regard. They conduct a field experiment on property tax collectors in Pakistan and find that performance-pay schemes can be a powerful device to improve tax administration and increase tax revenue. Different from the motivation of my study, their research focuses more on the trade-off between benefits from tax revenue and potential costs in social dissatisfaction and corruption.

[2] This paper reveals how the governments taxation respond to an adverse fiscal shock in the short-term. It complements Besley and Torsten (2009), who study the long-run response of states' taxation capacity to more drastic adverse shock such as external wars.

rate, which is the rate of tax that firms should legally pay. I find that the effective VAT rate increased on average by 8% due to the abolition of the agricultural tax. This implies that tax collection by county governments and tax authorities is not strictly constrained by access to information. I also show that this change is not driven by a change in reported value added: authorities are able to increase tax revenue by imposing higher effective tax rates on the reported tax base.[3]

Finally, this research examines several factors in an institution of multi-level governments that may affect the incentives to enforce VAT. The results show that the incentives could be attenuated if the county: (1) receives a lower proportion of total VAT revenues after sharing with prefectural and provincial governments; (2) has a broader VAT tax base; or, (3) has more abundant sources of revenue from land sales. The first finding confirms that, in an institution of multi-level government, tax-sharing system does shape the tax incentives of lower level of governments. The other two findings suggest that, due to tax competition and local politicians' personal interests originated from corruption (Chen, 2015b) or GDP-based promotion scheme in China, local governments are more liable to loose tax enforcement once they are more capable to mobilize fiscal revenue.

The striking contradiction between easy access to information and prevalent weak VAT enforcement in China offers a unique opportunity to study the role of political incentives in tax administration. *On the one hand*, VAT has long since been lauded for its advantage in terms of information revelation, as it leaves a paper trail along the supply chain (Pomeranz, 2013). In China, VAT is under the administration of the State Administration of Taxation (SAT) which aims to operate uniformly across the whole country and independently of subnational governments. With the help of the Nationwide VAT Special Invoices Online Cross-checking System, accessing relevant information is a simple procedure.[4] In addition, firms in the current study are large and formal and closely monitored by governments.[5] *While on the other*, VAT in China had not been well

[3]To understand this, define the taxation ratio as "tax payment/true tax base". The taxation ratio can be divided into two components: (1) report rate = reported tax base/true tax base; and, (2) effective tax rate = tax payment/reported tax base. Given the true tax base, tax revenue can be increased, either by cracking down on the under-reporting of tax base, or by raising the effective tax rate. Most recent empirical studies mainly focus on the first margin and highlight the role of information (Slemrod et al., 2001; Kleven et al., 2011; Pomeranz, 2012). Their typical approach is to identify the variation of reported tax base in field experiments, under the assumption that the true tax base does not change with the treatment. This paper focuses on the effective tax rate.

[4]The Nation-wide VAT Special Invoices Online Cross-checking System was set up in 2001. With this system, tax authorities can easily identify a huge variety of relevant information, including name of products or services in a transaction, quantity, value, and applicable VAT rate, and names of buyers and sellers, and so on.

[5]Normally, these big manufacturing firms are known as "above-scale industrial firms". Before 2006, a manufacturing firm in China is defined by the Bureau of Statistics as "above scale" if it is a state-owned enterprise or its sales volume is greater than five million Yuan (approximately $810,000 USD). From 2007, a state-owned enterprise with a sales volume lower than five million Yuan is not known as an above-scale firm. From January 2011, the threshold increased from five million Yuan to twenty million Yuan. The value-added of all above-scale firms accounts for about 90% of the total GDP of the manufacturing sector.

enforced. The empirical facts suggest that the effective VAT rate is not only commonly low across China, but also differs across regions.[6] Weak VAT enforcement, probably due to corruption at local level (Chen, 2015b), arising from local politicians' career concern, or driven by local tax competition, leaves great opportunities for subnational governments to raise VAT revenue by increasing the effective VAT rate.

This paper provides a direct contribution to the literature on the role of political incentives in tax enforcement. Most of the existing literature study the incentives of tax enforcement from perspectives of tax inspectors and tax administration. Besley and McLaren (1993) study the incentive role of wages for tax inspectors. Khwaja et al. (2014) study how performance-based payments could stimulate tax inspectors' efforts. Slemrod et al. (2001) find that audit threats deter tax misreporting behaviour. The role of information has been highlighed in literature, which find relevant parties, including product consumers in VAT (Pomeranz, 2013), third parties in income tax (Kleven et al., 2009, 2011; Kleven, 2014), final consumers (Naritomi, 2013), or workers (Kumler et al., 2013), can play an important role in information disclosure.[7] In addition, information can also explain the unique taxation structure in developing countries, and the formal financial sector is key to information access (Gordon and Li, 2009). This paper also aims to provide useful micro empirical evidence to recent studies that stress the interaction between economic development, tax systems and political institutions (Besley and Persson, 2009, 2011, 2013, 2014).

Additionally, the findings of this paper are consistent with a number of cross-country studies which claim that VAT administration is not effectively enforced in developing countries (Keen and Lockwood, 2010; Aizenman and Yothin, 2008). The VAT enforcement problem in China and its political determinants were initially documented by Chen (2015a, 2015b). Finally, this paper links to a huge literature on tax evasion, including the comprehensive review papers (Andreoni et al., 1998; Slemrod and Yitzhaki, 2002; Slemrod, 2007) and empirical evidence in China (Fisman and Wei, 2004; Cai and Liu, 2009).

The rest of the paper is organized as follows. Section 2 introduces the institutional background. Section 3 interprets the conceptual framework. Section 4 describes the datasets and measurements of the main variables. Section 5 introduces the empirical strategy and reports the main results. Section 6 studies the heterogeneity in incentives of VAT enforcement and the effects on total tax revenue and other government revenues. Section 7 discusses alternative explanations and conducts a variety of robustness checks. Section 8 concludes.

[6] For details, refer to Figure D.1 and D.2 in Appendix D.

[7] Of course, the role of the relevant party in information revelation cannot be exaggerated. For example, the effectiveness of third party information can be weakened if taxpayers can substitute misreporting to less verifiable margins (Carrillo, et al, 2014).

2 Institutional Background

2.1 VAT Administration

Two Separations in Government and Incentives for VAT Enforcement

The task of collecting tax cannot usually be efficiently delegated due to two separations in the public sector. *First*, there is a separation between governments and taxation administrators. The delegation of tax collecting by the former to the latter generally leads to a moral hazard problem. *Second*, there is a separation between national and sub-national governments. Conflicts in tax sharing or vertical tax competition are common between different levels of governments.

Figure 2.1: Four Players in VAT Administration in China

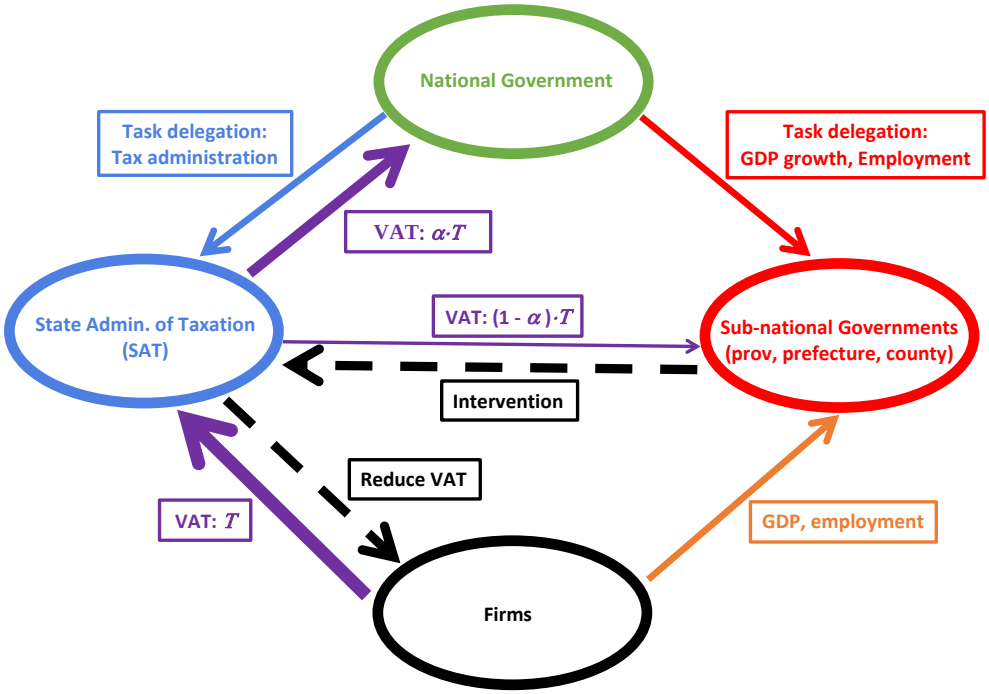


Figure 2.1 sketches the institutional background of VAT administration in China, with both types of separation in place. The figure shows that VAT administration in China is delegated to the State Administration of Taxation, which is under the direct leadership of the national government. The national government also generally delegates many executive tasks, such as GDP growth and employment, to sub-national governments. This leads to intense horizontal competition between sub-national governments across regions (Li and Zhou, 2005; Chen, Li, and Zhou, 2005). As a result, a race to the bottom in tax is inevitable. In addition, VAT is a sharing tax in China, with the national

government obtaining 75% (α in Figure 2.1) and sub-national governments, including provincial, prefectural, and county, taking 25%. This further intensifies horizontal tax competition. In an institutional setting like China, where local governments have no legislative right to set tax rates, the tax base, or to create a new tax, they have incentives to lean on the tax administration to lower the local effective VAT rate.[8]

Nevertheless, a weak incentive does not mean that local governments have no incentive at all to enforce VAT. Despite the low sharing ratio, VAT is still the second biggest source of tax revenue after business tax for subnational governments.[9] Also, VAT is the biggest source of tax revenue in about 1/5 of all counties (in 1/10 of all provinces).[10]

Subnational Government and VAT Administration

The central government controls the personnel appointments for the State Administration of Taxation (SAT) and its fund is allotted via a vertical administrative system. In spite of this, local governments can still exert enormous influence, either directly or indirectly, on VAT inspectors.

In the first place, the chief of a local office of the SAT is under the direct supervision of the party secretary in the same jurisdiction.[11]

In addition, there are several other indirect ways for subnational governments to intervene in VAT enforcement. *First*, the capability of the SAT in tax law enforcement is limited by local departments of Public Security, which are under tight control of local governments.[12] *Second*, the local government can help the SAT with issues such as obtaining land for office buildings, schooling for children, local hospitals for health care,

[8]Studies have shown that local governments in China offer preferential corporate income tax treatment in order to attract firms to invest and stay in their jurisdictions (for example, see Wu, 2009). Anecdotal evidences shows that VAT could also be used as a tool in tax competition.

[9]During the period of 2000 - 2007, VAT accounts for about 15% of the county government tax revenue (about 12% at province level), and business tax is about 25% at the county level (18% at province level).

[10]Roughly speaking, VAT in China is a tax on value-added in the manufacturing sector. Also, business tax is the one on turnover in the service sector. The third largest source of local tax revenue is corporate income tax on profits in both the manufacturing and service sectors, which accounts for about 7% of the county governments' tax revenue.

[11]China is under the governance of a hierarchical bureaucratic system, from central government at the top, followed by province, prefecture, county and village/town. At each level of government, there is an appointed secretary of the Communist Party who is the most powerful local official. At the same time, for each level of government, except village/town, there is a corresponding Nation Bureau of Taxation office (or the headquarters at the central government level). The chief of the local office of the State Administration of Taxation is appointed by the headquarters at the central level or by the office at the upper level of the State Administration of Taxation. However, in almost all cases the chief of the local office of State Administration of Taxation is also a party secretary of the local office. Within the system of the Communist Party of China, he/she is under the direct leadership of the party secretary in the same locality. For example, the chief/party secretary of a county office of the SAT is led by the county secretary of Communist Party. The latter can considerably affect the promotion and political career of the former.

[12]For example, in China, the SAT is entitled to crack down on fake invoices only with the help of the police. Therefore, the SAT can do nothing with a fake invoice if the local police are not willing to collaborate.

and so on. *Third*, it is openly known that the SAT receives a subsidy from the local government in order to improve its working conditions, or for any other reason. *Fourth*, tax administrators have their own dirty laundry. As a last resort, selective anti-corruption is a credible threat by local governments against a tax administrator once he/she refuses to cooperate.

To sum up, although the *de jure* VAT rates are legislated at a national level, local party secretaries and other local officials can influence the *de facto* tax rate of a firm.[13]

Ways to Lower the Effective VAT Rate

There are many ways to reduce the effective VAT rate. Here I introduce four possible cases. *First*, local governments can demand that the SAT turns a blind eye to firms using fake invoices, which are rampant across China.[14] In some extreme cases, local governments even subsidize and encourage local firms to use fake invoices.[15] *Second*, local governments can directly order the local office of the SAT to enforce a lower VAT rate on a firm, if the local officials are powerful enough.[16] *Third*, slack tax enforcement by local governments could be conducted through nominal delayed tax payment, so that the actual payment of a firm is smaller than its real tax liability. This delayed tax payment could be carried forward and not paid out unless demanded by governments. *Fourth*, although not quite as popular, some tax farmings are used in VAT collection. This makes room for flexible tax enforcement.[17]

[13] Chen (2015b) studies the role of the prefectural secretary of the Communist Party in selective enforcement on VAT across industries.

[14] Most of the VAT special invoices are not in fact fake. Actually, they are authentic invoices simply with a fake transaction or a fake recorded amount. There are many informal companies in China specialising in the sale of fake invoices. Figure D.4 in Appendix D shows records of my conversations with fake invoice dealers through text messages on the eligibility, price and sources of the VAT special invoices.

[15] Anecdotal evidence shows that this type of case is common in many regions. A case is currently being brought to the court in Jiang Xi province in which a company is being charged with issuing and selling fake VAT special invoices to other firms. The company under charge is said to be subsidized by the local government. For another example, the SAT announced the eight biggest cases of tax law violation in 2013. All of these cases were related to fake VAT special invoices. In every case, the tax evaded is above one hundred million RMB yuan (about $15 million USD). The highest effective tax rate a firm paid in one of these cases was only 0.11% (Xinhua Net, October 21, 2013. http://news.xinhuanet.com/2013-10/21/c_117804571.htm)

[16] This case is rare but not impossible. For example, anecdotal evidence reveals that in 2012, Samsung Electronics invested in a project worth £7b USD in Xi'an, the capital city of Shannxi Province. To attract the gigantic FDI, Mr. Zhao, the provincial secretary of the CPC at the time, agreed that a VAT rate of 11% could be applied to this project. Of course, disputes over the unlawful tax rate followed between the Shannxi government and the SAT. The central government had to step in, and finally Shannxi government won the case. It turned out to be a gentleman's agreement through private bargaining between the local government, the SAT and the central government.

[17] Tax farming is called "Bao shui" in Chinese. It means tax collection is contracted with tax payers who are requried to make tax payment on a basis of a pre-determined tax base, rather than the statutory tax base.

2.2 Agricultural Tax Reform

Agricultural taxation in China before the 2005-reform includes three types of taxes: *Agricultural Tax*, *Agricultural Specialty Tax* and *Livestock Tax*. These taxes are levied on various agricultural products. The tax-payer is an organisation or an individual who obtains revenue from the taxable agricultural products.[18] The agricultural taxation is a local tax and the revenue is mainly shared by county and prefectural governments.

The purpose of the central government in abolishing agricultural taxation was to relieve the burden of rural farmers and mitigate rural social unrest, which has been violent and rampant since the 1990s. To pave the way for the agricultural tax reform, and guarantee local government expenditure on education and government regular operation, the central government decided to subsidize local governments through an inter-governmental transfer/subsidy. The amount of subsidy is based on a formula and subsidy payments began in 2001.[19]

All agricultural taxes were abolished by the Standing Committee of the 10th National People's Congress from January 1st, 2006. Nevertheless, under the pressure and order of the central government, local governments nearly stopped collecting all these taxes in 2005, even before the law was formally passed by the national congress.[20] Therefore, the

[18] *The Agricultural Tax* is a type of tax levied on agricultural products, including grain, cotton, plant oil and sugar, among others. In legal terms, the Agricultural Tax is calculated based on the formula: *Agriculture Tax= Area of Land* $\times$ *Output per Unit of Land* $\times$ *Tax Rate.* . In practice, the output of crops per unit of land stems from a historical level in the early 1960s and did not change after that time, rather than using the actual output of current year. Therefore, the agriculture tax is literally more like a tax on land, rather than on agriculture products. After the 1990s, the agriculture tax was generally paid in cash based on the current price of crops rather than being paid in kind. The tax rate is determined by local governments, varying between 13% and 17%. The country average is 15.5%.

The Agricultural Specialty Tax is similar to the agriculture tax but levied on only some selected agriculture products, such as tobacco, tea, fruits, aquatic products, animal hides and furs, mushrooms, among others. The agriculture specialty tax is calculated based on the market value of the product, and the tax rate varies for different types of agriculture products. Similarly, *the Livestock Tax* is levied on livestock including horses, cows, sheep and goats and camels. The livestock tax is calculated either based on the market value of livestock or based on the number of herds.

[19] The factors used in the subsidy formula is information mainly relevant prior to the subsidy being made, including local expenditure on education, health, road construction, militia training, as well as the number of villages and households under the poverty line, and the number of students in primary and high school. In the whole country, the transfer from 2001 to 2005 respectively is (in Billion RMB Yuan): 8, 24.5, 30.5, 52.3, 65.4. 6. The formula for subsidy/transfer is as follows: subsidy = township transfer + village transfer + education fund transfer, where: township transfer = (village education expenditure + birth control expenditure + poverty alleviation expenditure + village road construction and repair expenditure + militia training expenditure + other funds expenditures + reduction in slaughter tax + reduction in agricultural specialty tax − increase in agricultural tax) $\times$ transfer coefficient. And the transfer coefficient = (ratio of agricultural tax revenue to total tax revenue / national-wide average ratio of agricultural tax revenue to total tax revenue $\times$ weight + ratio of basic public expenditure to total tax revenue / national-wide average ratio of basic public expenditure to total tax revenue $\times$ weight) $\times$ fiscal burden ratio of central government. And the village transfer = wage and compensation of village public servants + village poor household subsidy + administrative and office expenses. Readers can refer to Li and Xu (2006) for details.

[20] Although some local governments still collected agricultural taxes in 2005, it is very small and amounts only to 8.9% of collection in 2004. In empirical work, the author alternatively chose year 2006 as the

exact timing of the revenue loss was somewhat of a surprise for local governments, even though they had known since the early 2000s that the tax would be abolished sooner or later.

It should be noted that the agricultural tax is collected by the *Local Administration of Taxation*, and that VAT is administered by the *State Administration of Taxation*; two independent taxation bureaus. The former is under the direction of local governments, while the latter is led by the central government. The staff of the Local Administration of Taxation would not be re-assigned to the State Administration of Taxation following the abolition of the agricultural tax.

3 Conceptual Framework

Why does the abolition of agricultural tax serve as a good natural experiment to identify the impact on tax enforcement? What is the underlying economic mechanism? These questions may be answered by a model of optimal taxation with many types of tax available and taxation costs to be incurred.

The main mechanism of the model is the tension between benefits of public funds and taxation cost. The latter includes the administrative cost, deadweight loss, and local politicians' personal loss. The first two costs are common in the taxation theory. The third cost is emphasized in this paper. It is rooted from local politicians' personal interests due to corruption or career promotion, or because of jurisdictional tax competition under China's political institution. And it mainly comes from the fact that the likelihood of the promotion of local politicians is positively related to local GDP performance in China (Li and Zhou, 2005; Chen, Li, and Zhou, 2005; Xu, 2011), and higher tax may hurt investment and GDP growth.

In this section, I only sketch the model and its predictions relevant to my empirical work. Model details and rigorous technical proofs of predictions can be found in Appendix A.

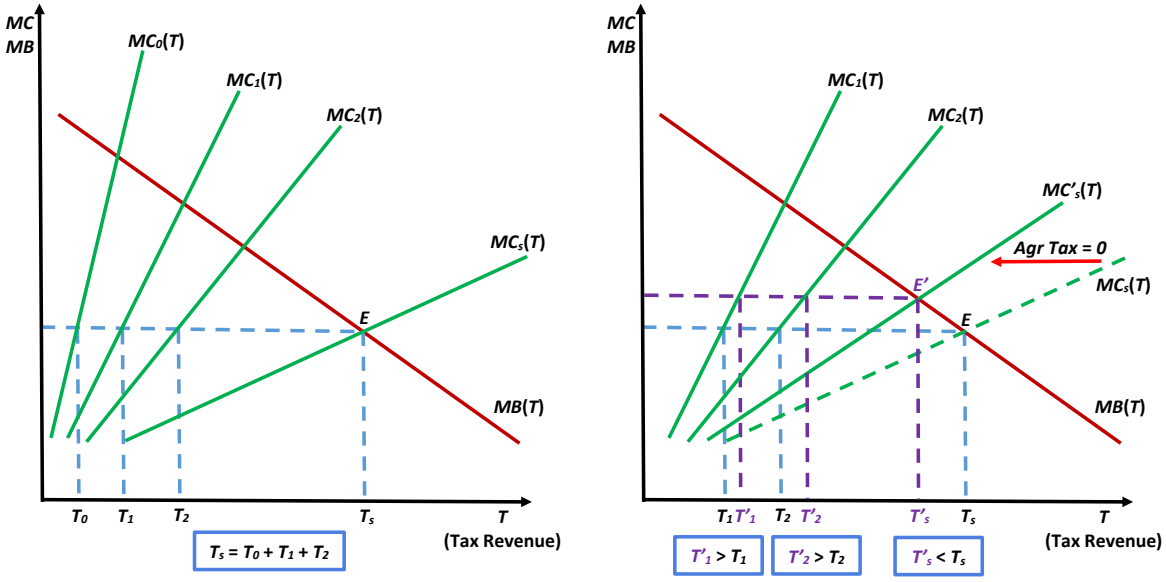
Suppose the public funds T can bring benefits $B(T)$ to local governments, with $B'(T) > 0$, and $B''(T) < 0$. Without loss of generality, I assume there are only three types of taxes to finance the funds, with each type of tax bearing a certain taxation cost denoted by $C_i(T_i)$, $i = 0, 1, 2$, with $C_i'(T_i) > 0$, and $C_i''(T_i) > 0$. Empirically, we can think of these three types of *taxes* respectively as agricultural tax, VAT, and land sales revenue. Of course, they can be generalized to other taxes or public revenues depending on the scenarios under study.

In this model, the marginal taxation costs $C_i'(T_i)$ are predominantly affected by three

reform year in addtion to the baseline regression with 2005 as reform year. The results in both cases are consistent.

types of costs: (1) the administrative costs, (2) the efficiency loss due to tax distortions, and, (3) the cost of tax base loss to local politicians, for instance, due to political promotion, horizontal tax competition between jurisdictions, or for other reasons. Appendix A presents details of how the political promotion and tax competition affect the taxation cost $C_i(T_i)$.

Figure 3.1: Effect of Abolition of Agriculture Tax



The optimal condition requires that the marginal costs of any tax should be equal to the marginal benefits of total public funds. That is,

$$B^{'}(T) = C_i^{'}(T_i) \qquad ,\forall i = 0, 1, 2 \tag{3.1}$$

The optimal condition (3.1) can be depicted by the left graph of Figure 3.1, where $MB(T)$ indicates the marginal benefit, and $MC_i(T)$ $(i = 0, 1, 2)$ are marginal cost of taxation respectively for three types of taxes. $MC_s(T)$ curve is the aggregate marginal cost of total public funds T. It is the horizontal aggregation of all the $MC_i(T)$ curves. The optimal total public funds T_s is determined by the intersection point E between $MB(T)$ and $MC_s(T)$ curve. Given $MC_s(T_s)$, the optimal amount of each tax T_i is equal to $MC_i^{-1}(MC_s(T_s))$, where $MC_i^{-1}(\cdot)$ is the inverse function of $MC_i(\cdot)$. That is, tax revenue T_1, T_2, and T_3 are set so that all marginal costs are equalized.

Suppose tax-0 (the agricultural tax) now has to abolished. What is its impact on the enforcement of other taxes and the mobilisation of other sources of public revenue? How can this impact be affected by other relevant factors. These questions can be answered by the following predictions based on the model.

The first question can be answered by Prediction 1 as below.

Prediction 1 (Abolition of Tax) *Suppose that tax-0 is abolished and no new type of tax can be introduced, then the government would strengthen enforcement on existing taxes and raise more revenue from other available sources.*

Crude graphical proof of prediction 1 can be demonstrated by the right panel of Figure 3.1. After the abolition of tax-0, the aggregate marginal cost curve $MC_s(T_s)$ should shift leftward, pushing the equilibrium from point E to point E'. This is because the marginal benefit of public funds rises due to revenue loss. Given rising marginal benefit, the government now has a stronger incentive to raise more revenue from existing taxes, even though it would be more costly to do so. In the end, both T_2 and T_3 increase. Generally, the total tax revenue T_s would decrease (or not change) if the marginal benefit curve $MB(T)$ is downward sloping (or vertical).

For the second question, I only consider three factors relevant to the empirical study: (1) tax sharing ratio of subnational governments;[21] (2) size of tax base; and, (3) marginal cost of other public funds. These three factors associate Prediction 2, 3, and 4, respectively.

Prediction 2 (Tax Sharing Ratio) *Suppose that tax-0 is abolished and tax-1's effective tax rate τ_1 increases, then τ_1 would rises by less if the local government had a lower sharing ratio in tax-1.*

The intuition of Prediction 2 is simple. A lower tax sharing ratio implies that local governments have to *collect* more total tax revenue from local firms in order to *obtain* a given amount of tax revenue. Therefore, a lower tax sharing ratio dis-incentivizes the local governments to enforce tax because the taxation cost for each Yuan of local governments' tax revenue would be greater.

Prediction 3 (Size of Tax Base) *Suppose that tax-0 is abolished and tax-1's effective tax rate τ_1 increases, then τ_1 would rises by less if the tax base y_1 is greater in tax-1.*

Prediction 3 is quite intuitive. What governments care about is total public funds. To raise a certain amount of revenue, they only need to increase the tax rate by a smaller amount if the tax base is larger.

Prediction 4 (Marginal Cost of Other Public Funds) *Suppose that tax-0 is abolished and tax-1's effective tax rate τ_1 increases, then τ_1 would rise by less if the marginal change of marginal cost (that is,$C_2''(T_2)$) of other funding sources is smaller.*

The rationale behind Prediction 4 is similar to that of Prediction 3. What governments care about is total public funds. They do not have to squeeze too much revenue from one

[21] Tax sharing is common in multi-level governments. For example, the VAT tax in China is a sharing tax. For 100 Yuan of VAT paid by a firm located in a county, 75 Yuan goes to the central government, the remaining 25 Yuan is divided between the provincial, prefectural and the county governments. Generally, the tax sharing ratio of a county government is determined at the subnational level. It therefore varies across provinces. Its geographic variation provides an opportunity to study the effect of the tax sharing ratio of a county government and its incentive to enforce tax. After all, county governments are at the grassroots level; their incentives can critically affect the tax enforcement. Readers can refer to Section 6.1 and Appendix B for details.

tax if they are actually able to mobilize revenue from other less costly sources.

4 Data and Main Variables of Interest

4.1 Data Description

This paper employs two datasets: (1) the Annual Survey of Industrial Production conducted by the National Bureau of Statistics of China (2000-2007); and, (2) the County Public Finance Statistics Yearbook of China (2000-2007).

The Annual Survey of Industrial Production includes all state-owned firms and non-state owned firms with annual sales of more than 5 million RMB yuan (approximately $800,000USD). I use data from 2000 to 2007. The number of firms increases from about 66,000 to 168,000 during the sample period after dropping bad observations.[22] Information on each firm includes a 4-digit industry code, ownership, county-level region code, value-added, sales revenue and tax payments including VAT, corporate income tax, business tax and other minor local taxes.

The County Public Finance Statistics Yearbook of China (2000 - 2007) includes government revenue and government expenditure at county and prefecture levels. In this paper, we are interested in the following variables: (1) tax revenue; (2) total revenue (including off-budget revenue); (3) total expenditure (including off-budget expenditure); (4) agricultural taxation revenue; (5) subsidies for agricultural taxation reform; (6) total population; and, (7) total GDP.

All the summary statistics of the main variables are reported in Table 1. It should be noted that Panel A reports the county-(2-digit) industry-year level data used for regressions. The data is collapsed from the firm level data of the Annual Survey of Industrial Production.

Table 1: Summary Statistics (Insert Here)

4.2 Measurement of Main Variables

4.2.1 Effective VAT Rate

I construct the effective VAT rate of a firm f in year t as:

$$Effective\ VAT\ rate_{f,t} = \frac{Payble\ VAT_{f,t}}{Value\ added_{f,t}} \tag{4.1}$$

[22] I drop three types of observations: (1) the observation with empty cell, and (2) the observation with non-positive value for capital stock, labor input, value-added and wage compensation, and (3) the observation with value-added less than the wage (even without adjustment in wage, value-added in many firms are also less than the wage compensation).

For notational convenience, I will ignore subscripts f and t hereafter for all relevant variables without causing confusion. The effective VAT rate defined in (4.1) can be further written as $\frac{\tilde{\tau}^s \cdot \tilde{S} - \tilde{\tau}^m \cdot \tilde{M}}{S-M}$, where $\tilde{S}$ and $\tilde{M}$ are the sales and intermediate inputs used to calculate the payable VAT. Due to tax evasion or other reasons, $\tilde{S}$ and $\tilde{M}$ may differ from S and M, which are used to calculate firms value-added. $\tilde{\tau}^s$ and $\tilde{\tau}^m$ are the tax rates actually applied for sales and intermediate inputs, respectively. They could differ from the statutory tax rates τ^s and τ^m. In the data, only $\tilde{\tau}^s \cdot \tilde{S}$ and $\tilde{\tau}^m \cdot \tilde{M}$ are observable.

It can be shown that the variation of the effective VAT rate may come from four sources: (1) $\frac{\tilde{\tau}^s \cdot \tilde{S} - \tilde{\tau}^m \cdot \tilde{M}}{\tau^s \cdot S - \tau^m \cdot M}$ due to tax enforcement or tax evasion; (2) statutory input VAT rate τ^m and output VAT rate τ^s ;[23] (3) input-output ratio M/S; (4) export-sales ratio E/S.[24]

Several caveats should be noted with the measurement of effective VAT rate in Expression (4.1). *First*, because only $\tau^s \cdot \tilde{S}$ and $\tau^m \cdot \tilde{M}$ are observable, we cannot distinguish the exact manner of tax evasion by manipulating tax rate (τ^m or τ^s) or by tax base ($\tilde{M}$ or $\tilde{S}$). *Second*, the variation of the effective VAT rate may come from the manipulation of reported value-added, that is, the denominator in Expression (4.1), rather than from tax evasion through the numerator. I will test the possibility of manipulation of reported value-added in Section 7.

4.2.2 Revenue Loss from Agriculture Tax Abolition

Although the agricultural taxation has already been a very small source of tax revenue in early 2000s, it was still, especially at the county level, an important part of local government tax revenue.[25] During 2000 - 2004, the agricultural taxation revenue on average accounted for around 12% of the county government tax revenue. Despite the subsidy for agricultural tax reform from the central government, the revenue loss incurred by the reform is still considerable to county governments.

[23] The statutory rate varies because certain sectors or activities are taxed at different rates, and firms differ in their input-output structure or to the extent to which they engage in activities subject to special tax treatment.

[24] To understand sources of variation in the effective VAT rate, I decompose it into two components:

$$Effective\ VAT\ rate = \frac{\tilde{\tau}^s \cdot \tilde{S} - \tilde{\tau}^m \cdot \tilde{M}}{\tau^s \cdot S - \tau^m \cdot M} \cdot \frac{\tau^s \cdot S - \tau^m \cdot M}{S-M} \tag{4.2}$$

where $\tau^s \cdot S - \tau^m \cdot M$ is the payable VAT calculated based strictly on the tax code.

The first component on the right-hand side of Expression (4.2), $\frac{\tilde{\tau}^s \cdot \tilde{S} - \tilde{\tau}^m \cdot \tilde{M}}{\tau^s \cdot S - \tau^m \cdot M}$, reflects the ratio of actual VAT and statutory VAT payments. The second component , $\frac{\tau^s \cdot S - \tau^m \cdot M}{S-M}$, can be re-written as $\tau^s + (\tau^s - \tau^m)(S/M - 1)^{-1}$, which implies that the variation in the statutory rates τ^s and τ^m, as well as the input-output ratio M/S, could potentially be sources of variation in the effective VAT rate. If the firm exports goods worth E, then τ^s in Expression (4.2) should be replaced with $\tau^s - \tau^e \cdot (E/S)$, where τ^e is the post-rebate statutory VAT rate for exports. In this case, the ratio of export to sales is an additional source of variation.

[25] For example, in 2001, the agriculture taxation was 28.6 billion RMB Yuan (approximately $4.67m USD) for the whole country, accounting for only 1.87% of total tax revenue of China (revenue of central government and local governments in total).

In this paper, I propose the following measurement of a revenue loss to each county c due to the abolition of agricultural taxation:[26]

$$Agr_c = \frac{(Agr\ Tax\ Revenue_{c,2000-2004} + Subsidy_{c,2000-2004})}{Total\ Tax\ Revenue_{c,2000-2004}} - \frac{Subsidy_{c,2005-2007}}{Total\ Tax\ Revenue_{c,2005-2007}} \tag{4.3}$$

On the right-hand side, $X_{c,2000-2004}$ ($X_{c,2005-2007}$) is the average of variable X in county c in a year between 2000 and 2004 (2005 and 2007). *Agr Tax Revenue* is the agricultural tax revenue. *Subsidy* is the formula-based central government transfer associated with the agricultural taxation reform and received at county level.[27] *Total Tax Revenue* is the total budgetary tax revenue, including VAT, corporate income tax, business tax and others, while off-budget revenue is not included.[28]

5 Empirical Strategy and Results

5.1 Motivating Facts

The nationwide abolition of agricultural tax in 2005 generated a variation of tax revenue loss from two dimensions. Over time, most counties suffered a net revenue loss. Across counties, the intensity of the revenue loss varied. The variation over time and across counties allows for the standard Diff-in-Diff method to identify the impact of agricultural tax abolition on tax enforcement.[29]

Given China's immense size and its geographical heterogeneity, it is obviously more plausible to use the variation within comparable locality rather than the variation across the whole country. In this paper, I choose to compare counties within the same prefecture. [30]

[26] Apparently, measurements of such a revenue loss are not unique. In Section 7, I will use alternative measurements for robustness checks.

[27] Refer to Section 2.2 for details of the subsidy. Figure D.5 in Appendix D shows that the subsidy began in 2002. No matter whether or not we include the subsidy into agriculture taxation as local government revenue, it is clear that local governments suffered a revenue loss in year 2005. We can also use a broader measurement of subsidy by also including subsidies due to abolition of agricultural specialty tax and reduction of agricultural tax rate, which started from year 2004. Due to lack of data in year 2007, the author does not use this measurement in calculating Agr_c in this paper. However, using the data from 2000 - 2006, a robustness check shows that Agr_c adopted in the paper is highly correlated with the Agr_c using the broader measurement of subsidy, and the main regression results do not significantly change with alternative measurement of Agr_c.

[28] In China, the local government can also raise revenue through other non-tax sources such as fees and local funds, or selling land and public assets such as state-owned enterprises.

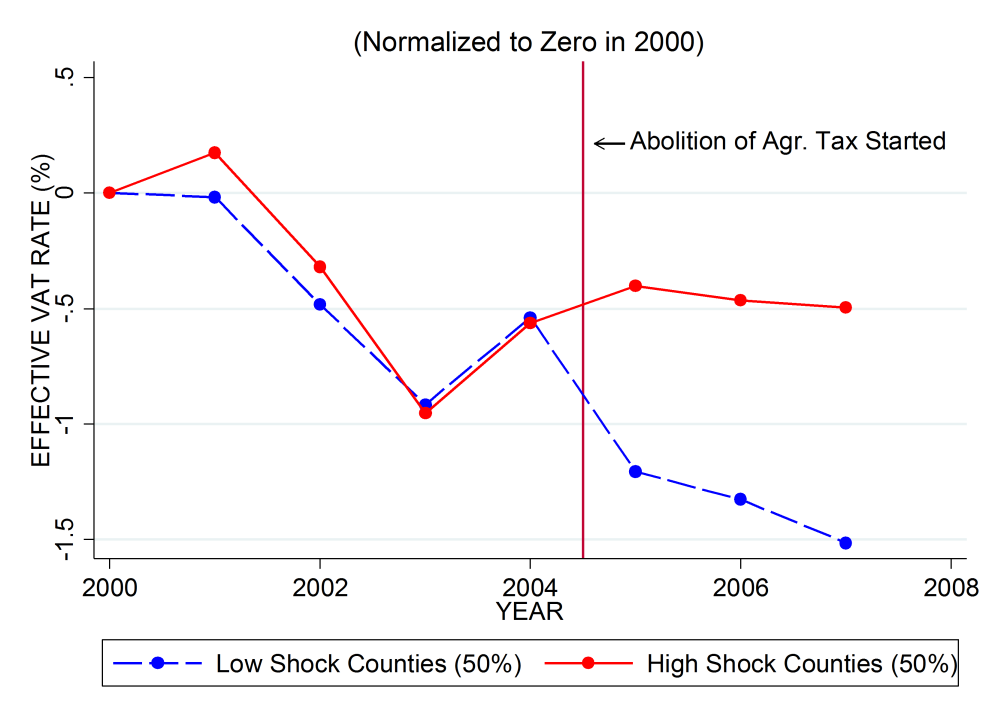
[29] Of course, the local government in China can also raise revenue through other non-tax sources such as fees and local funds, or selling the land as well as public assets such as state-owned enterprises.

[30] In China, there are about 3,000 counties and 300 prefectures in 31 provinces (including Beijing, Shanghai, Tianjin, and Chongqin as municipalities under direct administration of national government). On average, there are about 3 million people in each prefecture. Its area is approximately equal to a square of 180 kilometres. Nation-wide cross-county distribution of revenue loss is displayed in Figure D.5 (in histogram) and in Figure D.6 (in map). Within-prefecture cross-county distribution is shown in Figure D.7 (in histogram) and in Figure D.8 (in map with Aba Prefecture in Sichuan Province as an

The key assumption of the Diff-in-Diff method is that the control and treatment group should follow parallel trends before the treatment takes place. Also, the change of treatment group trend relative to that of the control group after treatment was thus identified as the treatment effect.

Firstly, I divide all counties into two groups based on a specified threshold in the intensity of revenue loss defined by Expression (3.3) after controlling for the prefecture-year fixed effects. I tentatively propose the median of as the threshold. Counties above (below) the median are treatment (control) group. To make a comparison, the effective VAT rates of both groups are normalized to zero in the year 2000. Then I plot the time profile for both groups. The figure suggests that a parallel trends assumption holds well before the abolition of agricultural tax. After 2005, something seems to have happened so that the parallel pattern broke down and there is a relative increase in effective VAT for the treatment group.[31] As a robustness check, I also set the 1st quartile of the intensity of the revenue loss as the threshold. The time profile of the effective VAT rate is plotted in Figure D.10 and D.11 in Appendix D.

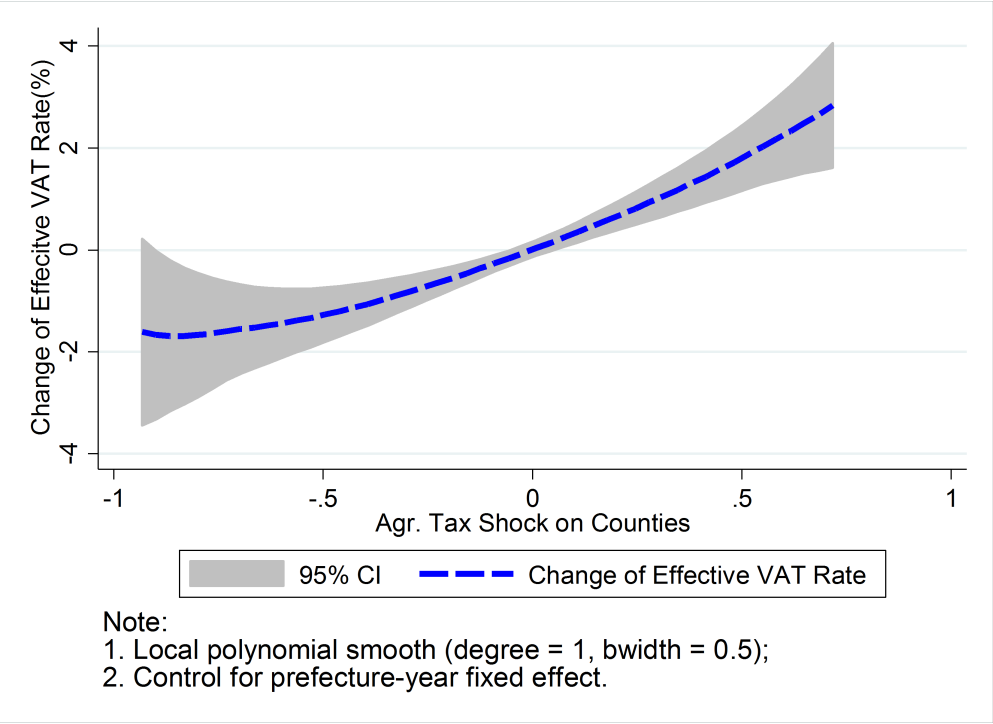
Figure 5.1: Effective VAT Rate over Time (Grouping by Median)



example).

[31] The time profile of the original data is plotted in Figure D.9 in Appendix D.

Figure 5.2: Change of Effective VAT Rate across Counties



If the change of the effective VAT rate after 2005 shown in Figure 5.1 is caused by the abolition of agricultural tax, then the magnitude of this change should be positively related to the intensity of revenue loss in each county. Therefore, we should expect a relative rise in the pre-post change in the effective VAT rate for counties suffering greater revenue loss.

This positive relationship is depicted in Figure 5.2, where the horizontal axis is the revenue loss defined by Expression (4.3), and the vertical axis is the pre-post change of the effective VAT rate for each county measured by the difference between the average effective VAT rate of period 2005 - 2007 and the average of 2000 - 2004.

5.2 Empirical Strategy

For suggestive purpose, Figure 5.1 and Figure 5.2 show that the variations of the effective VAT rate are correlated with revenue loss due to the agricultural tax abolition, both over time and across regions. In this section, I employ a standard Diff-in-Diff regression to identify the impact of tax revenue loss on the effective VAT rate.

The regression of Diff-in-Diff can be specified as below:

$$\tau_{c,p,t} = \alpha + \beta_c + \eta_{p,t} + \gamma \cdot Post_t + \theta \cdot Post_t \cdot Agr_c + \rho \cdot X_{c,t} + \epsilon_{c,p,t} \tag{5.1}$$

where the outcome variable $\tau_{c,p,t}$ is the effective VAT rate. The subscript c, p, t are county, prefecture, and year, respectively.

$Post_t$ is the dummy variable indicating the year pre and post the reform, taking value one for year 2005 and after. Agr_c, measured by Expression (4.3), is the revenue loss to county c.

Parameter of interest is θ, capturing the response of the effective VAT rate to the intensity of the tax revenue loss. County fixed effects β_c are included to absorb the county-specific characteristics. To guarantee that counties are comparable, the identification only exploits the very local within-prefecture variation, by controlling for prefecture-year fixed effects $\eta_{p,t}$, so that θ would only capture the county-year specific revenue loss within a prefecture.

$X_{c,t}$ is a vector of time-variant county characteristics, including median of firm size, capital-intensity, profitability, ownership dummy, and mobility dummy in county c and in year t.[32] They are included to avoid missing variables potentially correlated with $Post_t \cdot Agr_c$ and affecting $\tau_{c,p,t}$.

Identification assumes that there is no other county-year-specific shock correlated with the revenue loss due to agricultural tax abolition, or alternatively, counties should follow the parallel trend if tax abolition had not ever happened. To the author's knowledge, there is no other big reform in the same period correlated with the revenue loss and potentially has an impact on VAT rate. Of course, placebo and other various tests will be conducted in Section 7 to verify this assumption.

It should be noted that, even under parallel-trend assumption, different countries may respond differently to the shock of the same magnitude. Therefore, what θ captures is the Average Treatment Effect on the Treated (ATET), instead of the Average Treatment Effect on the Untreated (ATEU).

5.3 Empirical Results

Panel A of Table 2 reports the regression results of Equation (5.1). The positive coefficient 2.35 of $Agr \times Post$ shows that counties facing bigger revenue loss from agricultural tax abolition experienced a significant rise in the effective VAT rate, suggesting that the VAT enforcement there was strengthened after the reform.

What does the magnitude of the coefficient 2.35 mean? Suppose a 'treatment' county suffered 100 percentage points more revenue loss than a 'control' county. Then the coefficient 2.35 implies that the effective VAT rate rose by 2.35 percentage points more in the 'treatment' county compared to the 'control' county. On average, the gap in revenue loss between an average treatment county and control county is 30 percentage points in the sample. This means the effective VAT rate rose on average by 0.8 (=2.35 $\times$ 100 /

[32]Firm size is measured by LOG(sales), and LOG(Valued-added); capital-intensity by LOG(1+total assets/sales); profitability by total profit/sales; ownership dummy takes 1 for non-state-owned enterprises; mobility dummy takes 1 if the firm is in the mining industry. To avoid poor control, all these variables take the median within a county.

30) percentage points. Since the average effective VAT rate is 10 percentage points, this means it should have increased by 8 percent on average. Additionally, suppose there was no response in tax base, then the 8% increase in the effective VAT rate would translate into an 8% rise in VAT revenue.

Table 2: Agriculture Tax Reform and Effective VAT Rate (Insert Here)

6 Additional Results

6.1 Heterogeneous Effect on the Effective VAT Rate

In an institution where a local government's incentive to enforce tax codes is weak, even though they must strengthen VAT enforcement due to loss of tax revenue as shown in Section 5, their incentive could still be weakened by several factors.

In the conceptual framework in Section 4, I provided two predictions on the incentives of tax enforcement.

First, counties with a lower VAT sharing ratio may be more weakly incentivized to enforce VAT (Prediction 2).

Second, county governments may have a weaker incentive to raise the effective VAT rate if they have a stronger capacity to mobilize the revenue (Prediction 3 and Prediction 4. That is, they have either a greater VAT base or more capacity to obtain revenue from other sources such as land sales).

In this section, I will exploit the exogenous variation from the abolition of agricultural tax to test these two predictions.

6.1.1 VAT Sharing Ratio and Incentive of VAT Enforcement

VAT is a sharing tax between different levels of governments in China. The central government sharing ratio is 75%. This is the case across the whole country. However, the sharing ratio of provincial governments, prefectural governments and county governments differ across provinces. [33] The sharing ratio between subnational governments was set by subnational governments as early as 1994 when the historic tax reform was launched; it had not changed significantly since.[34] Some evidence suggests that a lower sharing ratio for county governments may attenuate their incentives to enforce VAT.[35] This is because the county government is at the grassroots level where tax enforcement is actually

[33]For example, county governments in Zhe Jiang province and Jiang Su province gain a high 70% of the province total VAT revenue (net of 75% taken by the central government), while those in Hei Long Jiang province only achieve 20%.

[34]Figure B.1 in Appendix B displays the cross-province variation of the VAT sharing ratio and its persistence over time.

[35]Table B.1 in Appendix reports the regression coefficients of effective VAT rate on VAT sharing ratio.

implemented, their low sharing ratios generate a weak incentive to enforce VAT, and the variation of the sharing ratio across regions could lead to dispersions in the effective VAT rate.

The abolition of agricultural tax provides us with an opportunity to study how the incentive of tax enforcement of local governments could be affected by VAT sharing ratio. Table 3 reports the heterogeneous effect of the abolition of agricultural tax on the effective VAT rate. Consistent with the results in Table 2, Table 3 shows that the abolition of agricultural tax leads to a rise in the effective VAT rate. In addition to this, it also suggests that the incentive of VAT enforcement could be weakened if county governments get a lower share of total VAT revenue. This can be seen from the negative coefficients on *Agr* × *Post* × (1 - *County Share*) in column (1), on *Agr* × *Post* × *Prefecture Share* in column (2), and on *Agr* × *Post* × *Province Share* in column (3).

I also run the horse-racing in column (4) by including both *Agr* × *Post* × *Prefecture Share* and *Agr* × *Post* × *Province Share* in regressions. The coefficients of both terms are still negative, but only *Agr* × *Post* × *Province Share* shows significance.[36]

It should be noted that, although the coefficients are significant, their magnitude is rather small. For example, in column (1), the coefficient is -0.23, implying that the rise of the effective VAT rate would be smaller by 0.00023 percentage points if the county sharing ratio decreased by 10 percentage points.

Table 3: VAT Sharing Ratio and Incentive of VAT Enforcement (Insert Here)

6.1.2 Revenue Mobilization Capacity and Incentive of VAT Enforcement

Some studies suggest that governments' incentive to build up taxation capacity could be weaker if they have stronger capacity to mobilize revenue from other sources (Jensen, 2011; Besley and Torsten, 2013). In this paper, I exploit the exogenous variation of the abolition of agricultural tax to study the similar mechanism for VAT enforcement.

Before doing this, I have two prior hypotheses in mind; that the incentives to enforce VAT can be affected by the following two factors. *First*, incentives are weaker in regions where county governments have a relatively broader VAT tax base. *Second*, incentives are weaker in regions where county governments have more abundant sources of revenue from land sales (a type of revenue from natural resources).

Empirically, these two factors can be measured in the following ways. *First*, the relative size of the VAT tax base is measured by the average of '*manufacturing GDP*/(*manufacturing and service GDP*)' between 2000 and 2004 for each county because VAT is mainly imposed on manufacturing firms in China. The agriculture sector is not included in the

[36]This may suggest the inefficiency loss along the delegation chain. The longer the distance between principal and agent, the weaker the incentive for the agent to take into account the principal's interests.

denominator because it can lead to a correlation with the revenue loss due to the agricultural tax abolition. *Second*, local governments' capacity to mobilize land sales revenue is measured respectively by two indices: '*Land sales revenue/local government revenue*', and '*land sales revenue/local GDP*'. These two indices also take the average between 2000 and 2004 for each county.

The results are reported in Table 4. Column (1) shows the heterogeneous effects associated with the relative size of VAT tax base. The significantly negative coefficient of $Agr \times Post \times VAT\ Base$ suggests that a greater VAT tax base weakens the incentive to enforce VAT. However, the magnitude is quite small, coefficient -0.21 implies that the effective VAT rate falls only by 0.021 percentage point if the ratio *manufacturing GDP/(manufacturing and service GDP)* increases by 10 percentage points.

Column (2) and (3) report the heterogeneous effects related to land sales revenue. The significantly negative coefficients of $Agr \times Post \times Land\ Sales/Rev$ and $Agr \times Post \times Land\ Sales/GDP$ suggest that the incentive to enforce VAT could be weakened by greater capacity to mobilize revenue from other sources. However, the magnitude is also very small: coefficients -0.26 (-1.77) suggest that the effective VAT rate falls by only 0.026 (0.00177) percentage points if the ratio *Land Sales*/Rev (*Land Sales/GDP*) increases by 10 percentage points.

Table 4: Resource Mobilization Capacity and Incentive of Tax Enforcement (Insert Here)

6.2 Effect on Total Tax Revenue and Other Revenues

In the end, we want to know the extent to which tax enforcement can offset revenue loss from the abolition of agricultural taxation. After all, the model in Section 3 predicts that the tax revenue net of agricultural tax should increase, while total tax revenue should drop, and local governments should also attempt to mobilize revenue from other sources; selling land could be an option in China.

Table 5 reports the effect on total tax revenue, other sources of revenue, and other measurements of county government revenue. All these revenue take logarithm. Panel A reports the effect on tax revenue. Column (1), (2), and (3) are the total tax revenue, total tax revenue net of agriculture and business tax revenue, respectively.[37] The results show that, while the total tax revenue insignificantly drops, the total tax revenue net of agricultural tax significantly increases. The coefficient of 0.39 implies that the total tax revenue net of agricultural tax increased on average by about 0.13% (= 0.39 × 10 / 3). The business tax revenue increased as well, by 8% (= 0.25 × 10 / 3).

[37]Business tax is a turnover tax imposed mainly on firms in the service sector. It is collected by the Local Administration of Taxation and is the largest source of local tax revenue. Due to lack of firm level data, this paper does not study the effect on firms' effective business tax rate.

Panel B extends the definition of government revenue by including revenue other than taxes. Column (4) is all-county government revenue including tax, transfer revenue and fund revenue. Column (5) is the budgetary revenue including taxes and transfer revenue. Column (6) is the land sales revenue, which has been a very important source of revenue for local governments since 2000. The results of all revenue and total budgetary revenue are insignificant. This is consistent with the effect on the total tax revenue in column (1). However, the land sales revenue significantly increased by 26% (= 0.78 × 10 / 3). [38]

The results in Table 5 show that county governments can raise the revenue, both by strengthening existing taxes and from other sources, such as land sales. The tax revenue loss due to the abolition of agricultural tax can be almost offset by the increase in other taxes. This suggests that the taxation capacity of county governments is not strictly constrained, even in the short run.

Table 5: Effect on Total Tax Revenue and Other Revenues (Insert Here)

7 Testing for Alternative Explanations and Robustness Checks

7.1 Placebo Test

The unbiased identification with Diff-in-Diff requires parallel trends. That is, THE outcome variable in different counties should follow the same time profile before the reform and show different trends after the reform. By specifying each year t before the reform with a dummy variable Yr_D_t, the following regression can be used to test the parallel-trend assumption needed for regression (5.1).

$$\begin{aligned}\tau_{c,p,t} = \alpha + \beta_c + \eta_{p,t} + \textstyle\sum_{t=2000}^{2003} \gamma_t \cdot Yr\_D_t + \sum_{t=2000}^{2003} \theta_t \cdot Yr\_D_t \cdot Agr_c \\ + \gamma \cdot Post_t + \theta \cdot Post_t \cdot Agr_c + \rho \cdot X_{c,t} + \epsilon_{c,p,t}\end{aligned} \tag{7.1}$$

Year 2004 is specified as the base-line year so its dummy are not included in the regression. The parallel-trend assumption requires that $\theta_t = 0$, for $t = 2000$, 2001, 2002, 2003.

Panel B of Table 2 reports the placebo test based on Equation (7.1), with year 2004 as the baseline year. The coefficients of year 2003, 2002, 2001 and 2000 are not significantly different from zero, confirming the parallel-trend assumption.

In addition, by running regression (7.2) as below, we can see the dynamics of difference in $\tau_{c,p,t}$ between counties suffering different levels of intensity of revenue loss due to the

[38] On average, land sales revenue account for 4% of all government revenue, or amount to 6% of total budgetary revenue.

agricultural tax abolition.

$$\tau_{c,p,t} = \alpha + \beta_c + \eta_{p,t} + \sum_{t=2000}^{2007} \gamma_t \cdot Yr\_D_t + \sum_{t=2000}^{2007} \theta_t \cdot Yr\_D_t \cdot Agr_c + \rho \cdot X_{c,t} + \epsilon_{c,p,t} \qquad (7.2)$$

In this setting, θ_t after year 2005 is expected to be positive when $\tau_{c,p,t}$ is the effective VAT rate. Similar to Equation (7.1), year 2004 is specified as the base-line year. The parallel-trend assumption now requires that $\theta_t = 0$ for any year prior 2005.

Figure 7.1: Dynamic Effect of AGR. Tax Reform on Effective VAT Rate

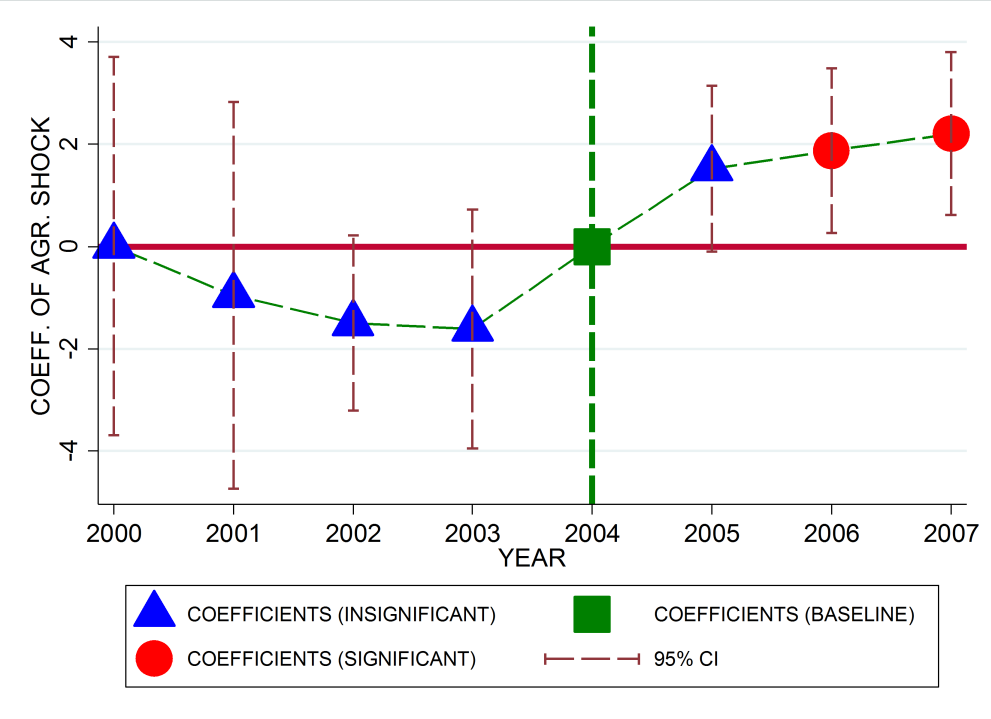


Figure 7.1 displays the dynamic effect of revenue loss of county governments on the effective VAT rate based on regression (7.2). In the graph, $\tau_{c,p,t}$ is normalized to zero for the county where the $Agr_c = 0$. The dots mark the $\tau_{c,p,t}$ for the county where $Agr_c = 1$, which means a 100 percentage points tax loss measured by Expression (4.3).[39] Year 2004 is the baseline year. We can see that before 2005, the effective VAT rate is not significantly different from the baseline year 2004, confirming the parallel-trend. In 2005, the year abolishing the agricultural tax, the effective VAT rate immediately increased in counties where $Agr_c = 1$. Although the increase is not significant at a 95% confidence level, it is at a 90% level. In 2006 and 2007, the effects continued to rise and they were all significant at a 95% confidence level.

[39] On average, the intensity of the revenue loss should be 30 percentage points, given the standard deviation of 15 percentage points within prefecture variation in the intensity of revenue loss shown in Figure D.8 in Appendix D.

7.2 Export-sales Ratio and Input-output Ratio

As discussed in Section 4.2.1, in addition to tax enforcement or tax evasion, a firm's effective VAT rate could also be affected by three other factors: (1) statutory input VAT and output VAT rate; (2) export-sales ratio (because of VAT refunds); and, (3) input-output ratio (due to intermediate input deduction).

By law, we know that there was no change in the statutory VAT rate during the sample periods in China. However, it is still possible that the abolition of agricultural tax affected the effective VAT rate through an export-sales ratio or input-output ratio. I explored these possibilities and report the results in Table 6, which show that there is no significant treatment effect either on *export/sales* or on *intermediate input/output*. Therefore, these two channels can be ruled out.

Table 6: Effect on Export-sales Ratio and Input-output Ratio (Insert Here)

7.3 Manipulation of the Value-added

From definition (4.1) of the effective VAT rate, we may suspect that the variation of effective VAT may come from the manipulation of value-added (the denominator) rather than from any change of VAT payment (the numerator). This suspicion is plausible given the fact that local governments in China have an incentive to over-report local GDP (aggregates of firms' value-added) for their political promotion.

I tried to test the validity of tax enforcement against the hypothesis of value-added manipulation in two ways. One way is to directly look at the effect of the abolition of agricultural tax on the value-added of firms. Another is to cross-validate the tax enforcement by studying the effect on the profit gap rate, which is thought to be related to corporate income tax evasion in China (Cai and Liu, 2009). The results suggest that the manipulation of value-added should be not a channel through which the effective VAT rate is changed.

7.3.1 Effect on the Value-added of firms

If the relative rise in the effective VAT rate in counties suffering more revenue loss is the result of manipulation of value-added, it must imply that the value-added of firms in these counties should be downwardly and relatively adjusted. However, this does not only contradict common knowledge that local officials have an incentive to rig up the value-added rather than the opposite. Additionally, it is inconsistent with the following facts that the value-added of firms in these counties did not actually significantly change.

Table 7 reports the effect of the abolition of agricultural tax on the value-added of firms. Panel A is based on Equation (5.1), and Panel B is on Equation (7.1). The

coefficient of '*Agr* × *Post*' shows the effect is not significant. The negative coefficients of '*Agr* × *Year 200x*' ($x = 3, 2, 1, 0$) imply that there was a relative rise in value-added after 2004 compared to year 2000 - 2003. Therefore, the evidence is not consistent with the *downward* manipulation of value-added following the abolition of agricultural tax.

Table 7: Effect on Firms' Value-added (Insert Here)

7.3.2 Effect on the Profit Gap Rate

The profit gap rate is the profit gap normalized the valued-added of a firm. Also, the profit gap of a firm measures the difference between its imputed profit and reported profit. Cai and Liu (2009) suggest that the profit gap is potentially related to tax evasion of corporate income tax in Chinese manufacturing firms (a larger profit gap implies more possible tax evasion).

Given the nature of the profit gap rate, I can cross-validate the effect on the effective VAT rate by studying the effect on the profit gap rate for two following reasons.[40] *First*, the conceptual framework in Section 3 predicts that the abolition of one type of tax would lead to the rise of tax enforcement on all available taxes. Therefore, the profit gap rate should decrease if the tax enforcement is strengthened. *Second*, since the effective VAT rate and the profit gap rate both have the value-added as the denominator, they should change in the same direction if it is the manipulation of value-added that leads to the variation of the effective VAT rate. Otherwise, they should move in opposing directions if it is the tax enforcement that results in these changes.[41]

Table 8 reports the baseline regression results based on Equation (5.1) and the placebo test based on Equation (7.1). The negative coefficients of '*Agr* × *Post*' suggest that firms in counties suffering more tax revenue loss are less likely to under-report profits following the abolition of agricultural tax. All the placebo tests confirm the parallel trends except the significantly positive coefficient of '*Agr* × *Year 2002*' in column (6).[42]

The dynamic effect based on Equation (7.2) is displayed in Figure C.3 in Appendix C.

Table 8: Effect on the Profit Gap Rate (Insert Here)

[40] Appendix C provides details on these reasons.

[41] Of course, I still cannot rule out the possibility that the value-added was manipulated and tax enforcement was strengthened simultaneously.

[42] To my knowledge, it is quite likely to be related to the corporate income tax reform in 2002 or another year-specific shock.

7.4 Robustness Checks

7.4.1 Change in Government Expenditure

So far I have only focused on the revenue side of the government budget. However, revenue and expenditure are related. The best outcome for the current study is that the expenditure of county governments remained rigid and did not change in response to the revenue loss: but was this the case?

To this end, I studied the variation of county government expenditure associated with the abolition of agricultural tax. The results are reported in Table 9. They show that the effects on ratios of government expenditure to revenue are positive (column (1) - (3)), while the effects on ratios of government expenditure to local GDP are not significant (column (4) - (5)). This implies that the county government expenditure was quite rigid while revenue fell. The expenditure did not make adjustments in response to revenue loss.

Table 9: Change in Government Expenditure (Insert Here)

7.4.2 Other Reforms

There are two major reforms during the sample period: (1) the corporate income tax reform in 2002;[43] and, (2) the transformation of VAT from production type to consumption type for eight industries in three north-east provinces in 2004.

To check whether these reforms could affect the conclusion of this paper, in column (1) of Table 10, I dropped all firms in the north-east provinces to ensure that the results would not be affected by the reform in 2004 of switching from production VAT to consumption. Also, in column (2), I dropped the firms that were set up after 2001 so that the firms that remained in the sample would not be affected by the corporate income tax reform of that year. The results in Table 10 show that the effect on effective VAT is still significant.

Table 10: Robustness Checks – Other Alternative Explanations (Insert Here)

7.4.3 Agriculture Goods Invoices

After the abolition of agriculture goods, firms which purchased agriculture products as intermediate inputs could no longer get the invoices for VAT deduction. This may increase the effective VAT rate of agriculture-related industries such as food processing, beverage, tobacco, textiles, leather and fur processing, among others. I drop all these industries from the sample, and the regression results are reported in column (3) of Table 10. The results show that the effect on effective VAT remains robust.

[43]This reform requires the State Administration of Taxation to be in charge of the corporate income tax for all firms established after 2002. Before the reform, the Local Administration of Taxation also collected corporate income tax of some firms according to firms' ownership and affiliation.

7.4.4 Entry of New Firms

China has experienced rapid economic growth with massive firm entry every year over the last two decades. To stimulate the investment, local governments generally promise preferential tax treatment to new firms. This could also affect the results of the regressions. To eliminate the channels through new firms, I drop the firms set up after 1999 so that no newly born firms are included in the sample period 2000 - 2007. The regression results are reported in column (4) of Table 10. They show that the effect on the effective VAT rate is still significant.

7.4.5 Change of Industrial Structure

The effective VAT rate varies across industries, even if tax enforcement strictly follows the tax code. If the industrial structure changes across regions, it may bias the regression results. To control for this, I run the regression at the county-year-(2-digit) industry level, and control for the county-(2-digit) industry fixed effect. The results are reported in column (5) of Table 10. Still, the effect on the effective VAT rate is robust.

7.4.6 Pilot Reform Regions

Before formally and legally abolishing the nation-wide agriculture tax, several pilot reforms were launched in a small number of regions. In 2000, Anhui province was chosen to start the pilot reform. In 2001, Jiangsu Province and a further 102 counties in other provinces joined the programme. In 2003, eleven provinces followed up, however; all pilot reforms were not compulsory in the sense that local governments were at their own discretion in determining the ways to reform and the magnitude to which they were to lower the agriculture taxation. Even within a province, different prefectures had their own views. Because the pilot programmes were not pervasive across the whole country and agriculture tax was not entirely abolished until 2005, the nationwide abolition in 2005 was still a considerable disturbance to these piloting regions.

Among all pilot reforms, two of the largest and earliest took place in Anhui province in 2000 and Jiangsu province in 2001.[44] To check whether the revenue loss measured with Expression (4.3) is still valid in pilot reform regions, I run regressions based on Equation (5.1), with counties in Anhui and Jiangsu provinces taking additional interaction between Agr_c and the *Post* dummy variable.[45]

[44] Figure D.12, D.13 and D.14 in Appendix D display the distribution of revenue loss in these two provinces. The revenue loss is still measured with Expression (4.3). The two graphs suggest that these still exist big within-prefecture variation in revenue loss despite these two provinces had launched the reform as least four years before the nationwide abolition in 2005. It suggests that the pilot regions still experienced different revenue loss across counties in 2005.

[45] Other sporadic pilot reforms are either small in scale, or very close to the year 2005. Their potential to invalidate the identification in this paper should be much weaker than those in Anhui and Jiangsu.

Table 11 reports the regression results. Column (1), (2), and (3) reports the regressions with Anhui province, Jiangsu province and the two provinces combined taking additional dummies respectively. The coefficients on $Agr \times Post \times Anhui\ Dummy$ and $Agr \times Post \times Jiangsu\ Dummy$ are close to zero and insignificant, suggesting that the pilot regions were not significantly different from other regions. In column (4), I drop observations from these two provinces and run the baseline regression based on Equation (5.1). Consistent with previous results, the coefficient on $Agr \times Post$ is still positive and significant at a 5% confidence level.

Table 11: Robustness Checks – Pilot Reforms (Insert Here)

8 Conclusion

Information and incentives are two pillars that support a well-functioning tax collection system for a country. Abundant theoretical analysis and empirical works have been conducted on the role of information. In contrast, the number of empirical studies on the incentives remains relatively small. Using firm-level data from the Annual Survey of Industrial Production (2000 - 2007), this paper exploits the nation-wide abolition of agricultural tax in China as a natural 'fiscal squeeze' experiment in which county governments were forced to raise taxation from firms to cover their rigid public expenditure.

The paper makes several contributions to our understanding of the extent and mechanisms of the role of incentives for low tax revenue in developing countries through the lens of China. *First*, the paper shows that a revenue loss on county governments in China can be largely and quickly offset by tougher tax enforcement. On average, the total tax revenue net of agricultural tax increased by 10% immediately in the year of agricultural tax abolition. This implies that the taxation capacity of developing countries may not be formidably binding, even in the short run.

Second, to understand whether incentives for tax enforcement could be a flexible margin to raise tax revenue, I particularly focus on the VAT. The paper has a novel finding in that tax enforcement can be made through raising the effective VAT rate. This implies that, given the information on the tax base (firms' value-added), there is still a room for tougher tax enforcement by imposing a greater effective tax rate. On average, the effective VAT rate increased by 8% immediately in the year of agricultural tax abolition and the increase lasted for at least three years. This finding is complementary to recently growing literature which focuses on the channel of raising tax, mainly through cracking down on the under-reporting of tax base.

Third, the incentives for VAT enforcement depend on several factors. The results

Therefore I do not intend to discuss them in the robustness check.

show that it can be weakened if the county: (1) receives a lower proportion of total VAT revenues after sharing with prefectural and provincial governments; (2) has a broader VAT tax base; or, (3) has more abundant sources of revenue from land sales.

Due to the limited data, the paper failed to answer several important questions. For example, through which channel was the effective VAT rate raised after the abolition of agricultural tax? Is it mainly through input VAT or output VAT? Or, through clamping down on fake invoice usage? Moreover, in addition to several factors studied in this paper, there should be many other factors that can constrain the incentives for tax enforcement. These questions are open for future research.

As Bird (2004) points out, "the best tax policy in the world is worth little if it cannot be implemented effectively". Despite the good nature of VAT in terms of information, its technological frontier cannot be fully achieved if tax administration is not well incentivized. Although the empirical study of this paper is based on the experience in China, it is not unique to the rest of the world, especially for countries where the task of collecting tax is not efficiently delegated. The inefficiency in delegation could be caused by the moral hazard of tax inspecting (Besley and McLaren, 1993), or the result of tax competition and tax sharing which is prevalent in fiscal federalism regimes. All results suggest that a weak incentive for tax enforcement can also be an important cause of low tax-to-GDP ratio in developing countries.

References

1. Allingham, M. G. and Sandmo, A. (1972) "Income Tax Evasion: A Theoretical Analysis", *Journal of Public Economics*, 1, pp. 323-338.

2. Aizenman, J. and Jinjarak, Y. (2008) "The collection efficiency of the Value Added Tax: Theory and international evidence", *Journal of International Trade and Economic Development*, 17(3), pp. 391-410.

3. Andreoni, J., Erard, B. and Feinstein, J. (1998) "Tax Compliance", *Journal of Economic Literature*, 36, pp. 818-860.

4. Bertrand, M., Duflo, E. and Mullainathan, S. (2004) "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1), pp. 249-275.

5. Besley, T. and McLaren, J. (1993) "Taxes and Bribery: The Role of Wage Incentives", *Economic Journal*, 103, pp. 119-141.

6. Besley, T. and Persson, T. (2009) "The Origins of State Capacity: Property Rights, Taxation, and Politics", *American Economic Review*, 99(4), pp. 1218-44.

7. Besley, T. and Persson, T. (2011) *Pillars of Prosperity: The Political Economics of Development Clusters*. USA: Princeton University Press.

8. Besley, T. and Persson, T. (2013) "Taxation and Development", in Auerbach, A., Chetty, R., Feldstein, M. and Saez, E. (eds.) *Handbook of Public Economics*, vol. 5, Elsevier: Amsterdam.

9. Besley, T. and Persson, T. (2014) "Why Do Developing Countries Tax So Little?" *Journal of Economic Perspectives*, 28(4), pp. 99-120.

10. Bird, R. (2004) "Administrative Dimensions of Tax Reform", *Asia-Pacific Tax Bulletin*, 10, pp. 134-150.

11. Cai, H. B. and Liu, Q. (2009) "Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms", *Economic Journal*, 119(537), pp. 764-795.

12. Carrillo, P., Pomeranz, D. and Khwaja, A. (2014) "Tax Me if You Can: Third-party Cross-Checks and Evasion Substitution", *Working paper*. USA: Havard Kennedy School, Harvard University.

13. Chen, S. X. G. (2015a) "VAT Dispersion and Aggregate Efficiency in China", *Working paper*. UK: *London School of Economics*.

14. Chen, S. X. G. (2015b) "Political Determinants of Selective Tax Enforcement: Evidence from China", *Working paper*. UK: *London School of Economics*.

15. Fisman, R. and Wei, S. J. (2004) "Tax Rates and Tax Evasion: Evidence from 'Missing Imports' in China", *Journal of Political Economy*, 112(2), pp. 471-496.

16. Gordon, R. and Wei, L. (2009) "Tax structures in developing countries: Many puzzles and a possible explanation", *Journal of Public Economics*, 93(7-8), pp. 855-866.

17. Jensen, A. (2011) "State-Building in Resource-Rich Economies", *Atlantic Journal of Economics*, 39, pp. 171-193.

18. Keen, M. and Lockwood, B. (2010) "The value added tax: Its causes and consequences", *Journal of Development Economics*, 92(2), pp. 138-151.

19. Khwaja, A. I., Olken, B. A. and Khan, A. (2014) "Property Tax Experiment in Punjab, Pakistan: Testing the Role of Wages, Incentives and Audit on Tax Inspectors' Behavior", *Working paper*. USA: Center for Internatinoal Development, Harvard University.

20. Kleven, H. J. (2014) "How Can Scandinavians Tax So Much?", *Journal of Economic Perspectives*, 28(4), pp. 77-98.

21. Kleven, H. J., Knudsen, M., Kreiner, C. T., Pedersen, S. and Saez, E. (2009) "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries", *NBER working papers* series no. 15218. USA: National Bureau of Economic Research.

22. Kleven, H. J., Knudsen, M., Kreiner, C. T., Pedersen, S. and Saez, E. (2011) "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark", *Econometrica*, 79, pp. 651-692.

23. Kumler, T. J., Verhoogen, E. A. and Frias, J. (2013) "Enlisting Workers in Monitoring Firms: Payroll Tax Compliance in Mexico", *NBER working paper* series no. 19385. USA: National Bureau of Economic Research.

24. Li, P. and Xu, H. (2006) *Graphic Illustration of Inter-governmental Relationships in China* (*Zhong Guo Zheng Fu Jian Cai Zheng Guan Xi Tu Jie*). China: China Public Finance and Economy Publishing House.

25. Naritomi, J., (2013) "Consumers as Tax Auditors", *Working paper*, USA: Harvard University.

26. Slemrod J. (2007) "Cheating Ourselves: The Economics of Tax Evasion", *Journal of Economic Perspectives*, 21(1), pp. 25-48.

27. Slemrod, J., Blumenthal, M. and Christian, C.W. (2001) "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota", *Journal of Public Economics*, 79(3), pp. 455-483.

28. Wu, W. F., Wu, C. F., and Rui, M. (2009) "Government Background of Top Leaders in China's Listed Companies and Tax Preferential Policy", *Management World*, 3, pp. 134-142. (In Chinese)

29. Xu, C. G. (2011) "The Fundamental Institutions of China's Reforms and Development", *Journal of Economic Literature*, 49(4), pp. 1076-1151.

Table 1: Summary Statistics

	Panel A: Annual Survey of Industrial Production in China (2000 - 2007)							
	Sample Size	Mean	St. Dev.	Min	Max	p10	p50	p90
Effective VAT Rate[1]	139,287	10.34	9.66	-153.22	697.38	0.98	8.74	20.87
Profit Gap Rate[2]	139,287	34.67	42.85	-2783.55	2813.98	-12.3	40.19	75.92
Debt Ratio[3]	139,287	55.77	28.94	-120.02	1585.95	22.19	55.28	86.37
CapInt[4]	139,287	0.18	0.21	0	1.24	0.01	0.13	0.32
Ownership[5]	139,287	0.89	0.18	0	1	0.7	0.95	1
Profitability[6]	139,287	0.09	0.65	-16.8	176.27	0.03	0.06	0.14
Firm Age[7]	139,287	6.36	4.71	0	95	3	6	9
LOG(V-added)[8]	139,287	9.05	0.69	3.76	16.31	8.33	8.96	9.85
LOG(Sales)[9]	139,287	10.1	0.7	3.65	16.4	9.36	10.04	10.9
Input - output Ratio[10]	139,287	0.76	0.17	0	21.3	0.69	0.76	0.84
Export - Sales Ratio[11]	139,287	0.02	0.13	0	10	0	0	0
	Panel B: County Public Finance Statistics (2000 - 2007)							
	Sample Size	Mean	St. Dev.	Min	Max	p10	p50	p90
Agr[12]	19,950	0.25	0.23	-0.81	1.21	0.01	0.22	0.56
LOG(GDP)[13]	8,898	3.47	0.94	0	5.78	2.2	3.61	4.52

Notes:
1. "Effective VAT Rate" is defined as "Payable VAT / value-added × 100".
2. "Profit Gap Rate" is measured by "profit gap / value-added × 100", where "profit gap" is the "imputed profit" minus the "reported profit". See Section 5.1 for more details.
3. "Debt Ratio" is measured by "total liability/ total assets × 100".
4. "CapInt" is capital-intensity measured by "capital / sales". Capital is measured by "total asset" in Panel A, and measured by "net fixed asset / sales" in Panel B. The measurements are different due to lack of consistent data.
5. "Ownership" is dummy variable which is equal to 1 if the firm is not a state-owned enterprise.
6. "Profitability" is measured by "profit / value-added", indicating the degree of competition of 2-digit industry.
7. "Firm Age" is the years since the firm was set up.
8. "LOG(V-added)" is the logarithm of a firm's value-added.
9. "LOG(Sales)" is the logarithm of a firm's sales volume.
10. "Input-output Ratio" is defined as "intermediate inputs / total value of products and services".
11. "Export-sales Ratio" is defined as "export / sales volume".
12. "Agr" is the net tax revenue loss due to the abolition of agricultural tax. Refer to Expression (4.2) for its definition.
13. "LOG(GDP)" is the logarithm of a county's GDP per capita.

Table 2: Abolition of Agriculture Tax and Effective VAT Rate

	Dependent Variable: Effective VAT Rate					
	Panel A: Baseline	Panel B: Placebo Test				
	(1)	(2)	(3)	(4)	(5)	(6)
Agr × Post	2.35	2.22	2.11	2.37	2.41	1.78
	(0.82)***	(0.83)***	(0.84)**	(0.78)***	(0.78)***	(0.67)***
Agr × Year 2003		-1.05				-1.52
		(1.02)				(1.18)
Agr × Year 2002			-0.84			-1.18
			(0.74)			(0.87)
Agr × Year 2001				0.09		-0.57
				(1.55)		(1.90)
Agr × Year 2000					0.57	-0.17
					(1.48)	(1.85)
	Controls: Year Dummy					
Post	0.00	-0.01	0.01	0.00	0.00	0.02
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)
Year 2003		-0.06				-0.06
		(0.04)				(0.06)
Year 2002			0.03			0.04
			(0.04)			(0.030
Year 2001				0.05		0.03
				(0.04)		(0.06)
Year 2000					0.02	0.01
					(0.04)	(0.07)
Controls	*LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*					
	Fixed Effects					
Prefecture-Year FE	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES
Sample size	17,968	17,968	17,968	17,968	17,968	17,968
R^2	0.44	0.44	0.44	0.44	0.44	0.44
Clustering	Prefecture level (362 groups)					

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 3: VAT Sharing Ratio and Incentives for Tax Enforcement

	Dependent Variable: Effective VAT Rate			
	(1)	(2)	(3)	(4)
Agr × Post	2.68	2.62	2.58	2.68
	(0.84)***	(0.85)***	(0.83)***	(0.85)***
Agr × Post × (1 - County Share) / 100	-0.23			
	(0.06)***			
Agr × Post × Prefecture Share / 100		-0.29		-0.16
		(0.09)***		(0.10)
Agr × Post × Province Share / 100			-0.48	-0.35
			(0.15)***	(0.17)**
	Post and Other Interaction Terms			
Post	0.03	0.01	0.01	0.03
	(0.02)	(0.03)	(0.03)	(0.02)
Post × (1 - County Share) / 100	0.99			
	(1.58)			
Post × Prefecture Share / 100		1.88		0.57
		(2.9)		(3.37)
Post × Province Share / 100			2.90	1.73
			(3.39)	(3.93)
Controls	*LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*			
	Fixed Effects			
Prefecture-Year FE	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Sample size	17,968	17,968	17,968	17,968
R^2	0.44	0.44	0.44	0.44
Clustering	Prefecture level(362 groups)			

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 4: Resource Mobilization Capacity and Incentives for Tax Enforcement

	Dependent Variable: Effective VAT Rate		
	(1)	(2)	(3)
Agr × Post	2.56	2.43	2.50
	(0.85)***	(0.84)***	(0.84)***
Agr × Post × VAT Base /100	-0.21		
	(0.07)***		
Agr × Post × Land Sales / GDP / 100		-0.26	
		(0.11)***	
Agr × Post × Land Sales / Rev / 10000			-1.77
			(0.72)**
	Post and Other Interaction Terms		
Post	0.01	0.06	0.05
	(0.06)	(0.05)	(0.05)
Post × VAT Base / 100	2.22		
	(3.30)		
Post × Land Sales / GDP / 100		-9.99	
		(9.47)	
Post × Land Sales / Rev / 10000			-37.04
			(46.02)
Controls	*LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*		
	Fixed Effects		
Prefecture-Year FE	YES	YES	YES
County FE	YES	YES	YES
Sample size	17,828	17,872	17,944
R^2	0.44	0.44	0.44
Clustering	Prefecture level(362 groups)		

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 5: Effect on Total Tax Revenue and Other Revenues

	Dependent Variable: LOG(X)					
	Panel A: Tax Revenue			Panel B: Other Revenue Included		
	(1)	(2)	(3)	(4)	(5)	(6)
	Total Tax	Tax Net Agr	Business Tax	All Rev.	Bgt Rev.	Land Rev
Agr × Post	-0.05	0.39	0.25	0.01	-0.04	0.78
	(0.07)	(0.05)***	(0.07)***	(0.07)	(0.07)	(0.36)**
Post	0.01	0	-0.03	0.48	0.53	0.28
	(0.01)	(0.01)	(0.01)***	(0.02)***	(0.02)***	(0.04)***
Controls	*LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*					
	Fixed Effects					
Prefecture-Year FE	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES
Sample size	17,895	17,895	17,895	15,147	17,887	10,020
R^2	0.9	0.91	0.86	0.94	0.93	0.64
Clustering	Prefecture level (362 groups)					

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 6: Effect on Export-sales Ratio and Input-output Ratio

	Dependent Variable	
	(1)	(2)
	Intermediate Input / Output	Export / Sales
Agr × Post	-0.02	-0.99
	(0.84)	(1.26)
Post	-1.02	-2.37
	(0.22)***	(0.91)***
Controls	*LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*	
	Fixed Effects	
Prefecture-Year FE	YES	YES
County FE	YES	YES
Sample size	17,968	17,968
R^2	0.69	0.17
Clustering	Prefecture level (362 groups)	

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 7: Effect on Firms' Value-added

	Dependent Variable: LOG(Value-added)					
	Panel A: Baseline	Panel B: Placebo Test				
	(1)	(2)	(3)	(4)	(5)	(6)
Agr × Post	0.05	0.04	0.03	0.08	0.04	-0.05
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Agr × Year 2003		-0.16				-0.27
		(0.09)*				(0.12)**
Agr × Year 2002			-0.11			-0.2
			(0.09)			(0.12)*
Agr × Year 2001				0.15		-0.02
				(0.11)		(0.14)
Agr × Year 2000					-0.16	-0.27
					(0.11)	(0.12)**
	Controls: Year Dummy					
Post	-0.01	0.00	-0.01	-0.01	-0.01	0.00
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)*
Year 2003		0.02				0.03
		(0.00)***				(0.00)***
Year 2002			-0.01			0.00
			(0.00)***			(0.00)
Year 2001				0.01		0.02
				(0.00)***		(0.00)***
Year 2000					0.01	0.02
					(0.00)***	(0.00)***
Controls	*LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*					
	Fixed Effects					
Prefecture-Year FE	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES
Sample size	17,950	17,950	17,950	17,950	17,950	17,950
R^2	0.84	0.84	0.84	0.84	0.84	0.84
Clustering	Prefecture level (362 groups)					

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 8: Effect on the Profit Gap Rate

	Dependent Variable: Profit Gap Rate					
	Panel A: Baseline	Panel B: Placebo Test				
	(1)	(2)	(3)	(4)	(5)	(6)
Agr × Post	-13.68	-13.54	-12.33	-14.17	-12.03	-7.19
	(4.27)***	(4.45)***	(4.73)***	(4.43)***	(3.28)***	(2.70)***
Agr × Year 2003		1.41				9.51
		-6.09				-6.42
Agr × Year 2002			4.82			10.23
			-4.23			(4.29)**
Agr × Year 2001				-4.63		4.88
				-11.56		-11.45
Agr × Year 2000					17.54	23.17
					-18.26	-19.86
	Controls: Year Dummy					
Post	-0.12	-0.12	-0.19	-0.15	-0.15	-0.39
	(0.20)	(0.20)	(0.22)	(0.20)	(0.22)	(0.39)
Year 2003		-0.01				-0.01
		(0.20)				(0.40)
Year 2002			-0.13			-0.32
			(0.25)			(0.36)
Year 2001				-0.29		-0.24
				(0.23)		(0.44)
Year 2000					-0.1	-0.11
					(0.31)	(0.49)
Controls	*LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*					
	Fixed Effects					
Prefecture-Year FE	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES
Sample size	17,968	17,968	17,968	17,968	17,968	17,968
R^2	0.45	0.45	0.45	0.45	0.45	0.45
Clustering	Prefecture level (362 groups)					

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 9: Change in Government Expenditure

	Dependent Variable: LOG(1 + X / GDP)				
	(1)	(2)	(3)	(4)	(5)
	BGT EXP.	ALL EXP.	TAX REV.	BGT REV.	ALL REV.
Agr × Post	-2.21	-0.07	-5.47	-4.66	-4.46
	(2.88)	(3.02)	(2.61)**	(2.78)*	(2.52)*
Post	1.13	0.93	0.74	0.72	0.6
	(0.31)***	(0.26)***	(0.27)***	(0.26)***	(0.23)***
	Fixed Effects				
Prefecture-Year FE	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Sample size	17,741	12,694	17,741	17,740	15,023
R^2	0.79	0.73	0.74	0.73	0.65
Clustering	Prefecture level (362 groups)				

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 10: Robustness Checks – Other Alternative Explanations

	Dependent Variable: Effective VAT Rate				
	(1)	(2)	(3)	(4)	(5)
	Drop NE Provinces	Drop Firms > 2001	Drop Agr. Related Industries	Firm Entry (Drop > 1999)	Control for Industry FE
Agr × Post	2.82	2.46	2.39	1.9	1.31
	(0.87)***	(0.82)***	(0.84)***	(0.92)**	(0.48)***
Post	0.00	0.01	0.01	0.00	0.02
	(0.03)	(0.03)	-0.04	(0.04)	(0.04)
Controls	*LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*				NO
	Fixed Effects				
Prefecture-Year FE	YES	YES	YES	YES	NO
County FE	YES	YES	YES	YES	NO
Pref-Indu-Year FE	NO	NO	NO	NO	YES
County-Indu FE	NO	NO	NO	NO	YES
Sample size	16,228	17,625	17,356	17,264	11,938
R^2	0.43	0.44	0.47	0.45	0.17
Clustering	Prefecture (362 groups)	Prefecture (362 groups)	Prefecture (362 groups)	Prefecture (362 groups)	Prov-Indu (155 groups)

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups) in column (1) - (4), at the province-(1-digit)industry level in column (5). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.
2. The Pref-Indu-Year FE refers to Prefecture-(1-digit)Industry-Year fixed effect. The County-Indu FE is County-(2-digit)Industry fixed effect.

Table 11: Robustness Checks – Pilot Reforms

	Dependent Variable: Effective VAT Rate			
	(1)	(2)	(3)	(4)
Agr × Post	1.84	1.8	1.84	1.76
	(0.63)***	(0.62)***	(0.63)***	(0.68)**
Post	0.01	0.01	0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.03)
Agr × Post × Anhui Dummy	0.00	NO	0.00	NO
	(0.00)		(0.00)	
Agr × Post × Jiangsu Dummy	NO	0.00	0.00	NO
		(0.00)	(0.00)	
Controls	*LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*			
	Fixed Effects			
Prefecture-Year FE	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Sample size	17,950	17,950	17,950	16,671
R^2	0.5	0.5	0.5	0.5
Clustering	Prefecture level			
	(362 groups)	(362 groups)	(362 groups)	(332 groups)

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups in column (1) - (3), 332 groups in column (4)). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

A Appendix: Conceptual Framework

A.1 Basic Model

This section is an appendix to the conceptual framework in Section 3.

In this section I characterize the optimal taxation with many types of tax available and when taxation costs have to be incurred. Suppose public funds T can bring benefit $B(T)$, with $B'(T) > 0$, and $B''(T) < 0$. Without loss of generality, I assume there are only three types of taxes to finance the funds, with each type of tax bearing certain taxation costs denoted by $C_i(T_i)$, $i = 0, 1, 2$, with $C_i'(T_i) > 0$, and $C_i''(T_i) > 0$.

In this model, the marginal taxation cost $C_i'(T_i)$ is mainly affected by two types of costs: (1) the administrative costs, (2) the efficiency loss due to tax distortions, and, (3) the cost of tax base loss to local politicians in the presence of horizontal tax competition between jurisdictions. The third type of cost mainly comes from the fact that the likelihood of promotion of local politicians is positively related to local GDP performance.[1]

The problem facing the local government can be simply described by the following maximization problem:

$$max_{\{\tau_i\}_{i=0}^2} \qquad B(T) - \sum_{i=0}^{2} C_i(\tau_i \cdot y_i)$$

where $T = \sum_{i=0}^{2} T_i$, and $T_i = \tau_i \cdot y_i$, and y_i is the tax base, which responds to tax rate τ_i. Suppose taxation is on the good side of the Laffer curve, therefore $\partial T_i / \partial \tau_i > 0$.

The optimal condition requires that the marginal cost of any tax should be equal to the marginal benefit of total public funds. That is,

$$B'(T) = C_i'(T_i) \qquad , \forall i = 0, 1, 2 \tag{A.1}$$

[1] Suppose $C_i(T_i) = c_i(\tau_i) \cdot y_i$-$\beta \cdot p_i(y_i)$, where $c_i(\tau_i)$ is the administrative cost of tax collection and efficiency loss associated with the tax rate τ_i for one unit of tax base. β is the utility gain from political promotion of local politicians. $p_i(y_i)$ is the probability getting promoted, and assume $p_i'(y_i) > 0$ because promotion is positivey related to local GDP. The marginal taxation cost then can be written as: $C_i'(T_i) = c_i'(\tau_i)\left(1 - \frac{\tilde{p}_i(\tau_i)/\varepsilon^c_{\tau,i} - 1}{1/|\varepsilon^y_{\tau,i}| - 1}\right)$, where $\varepsilon^y_{\tau,i} = \frac{\tau_i}{y_i}\frac{dy_i}{d\tau_i} < 0$, and $\varepsilon^c_{\tau,i} = \frac{\tau_i}{c_i}\frac{dc_i}{d\tau_i} > 0$, $\tilde{p}_i(\tau_i) = 1 - \beta \cdot p_i'(y_i)/c_i(\tau_i)$. Obviously, $C_i'(T_i)$ is not only affected by $c_i'(\tau_i)$ and $\varepsilon^c_{\tau,i}$, but also by $\varepsilon^y_{\tau,i}$. The latter is related to the tax base loss caused by imposing a higher tax rate. Its underlying determinant is the degree of tax competition between jurisdictions.

Suppose $-1 < \varepsilon^y_{\tau,i} < 0$, then $\partial C_i'(T_i)/\partial \left|\varepsilon^y_{\tau,i}\right| > (=, <) 0 \Longleftrightarrow 1 - \beta \cdot p_i'(y_i)/c_i(\tau_i) < (=, >) \varepsilon^c_{\tau,i}$. That means, given the taxation costs $c_i(\tau_i)$ and $\varepsilon^c_{\tau,i}$, the greater is the utility from political promotion (bigger β) and the more sensitive is likelihood of promotion to GDP performance (bigger $p_i'(y_i)$), the more likely that a greater intensity of tax competition $\left|\varepsilon^y_{\tau,i}\right|$ leads to a bigger marginal taxation cost $C_i'(T_i)$.

A.2 Comparative Statics

Abolition of Tax

Now suppose that tax-0 is abolished and no new type of tax can be introduced. Curve $MC_1(T)$ and $MC_2(T)$ would not move, but the aggregate marginal taxation cost curve $MC(T)$ should shift leftwards to the position of $MC^{'}(T)$. This drives up the taxation cost and the optimal total public funds drop from T_s to $T_s^{'}$, while the tax revenue from both remaining available taxes go up, from T_1 to $T_1^{'}$, and from T_2 to $T_2^{'}$, respectively. Given the assumption that $\partial T_i/\partial \tau_i > 0$, we know both τ_1 and τ_2 must increase to sustain the public funds. Government can raise T_i in two ways as a response: (1) increase the effective tax rate; (2) crack down on tax evasion and tax base under-reporting.

The prediction can be summarized by the following statement.

Prediction 1. *Suppose that tax-0 is abolished and no new type of tax can be introduced, then the government would strengthen enforcement on existing taxes and raise more revenue from other available sources.*

Proof: To make comparative static analysis, I denote the marginal taxation cost as $C_i^{'}(T_i;\theta_i)$, $i = 0,\ 1,\ 2$, where θ_i is a parameter governing the marginal taxation cost, and it satisfies $\partial C_i^{'}(T_i;\theta_i)/\partial\theta_i > 0$. For tax-0, we additionally assume that $\lim_{\theta_0\to+\infty} C_0^{'}(T_0;\theta_0) = +\infty$. Under these assumptions, the abolition of agricultural tax is equivalent to raising θ_0 from a finite number to positive infinity (that is, collecting agricultural tax will be extremely costly). To prove prediction 1, we just need to show that $\partial T_1/\partial\theta_0 > 0$ and $\partial T_2/\partial\theta_0 > 0$. Here we go.

The optimal conditions (A.1) define a simultaneous equations systems in three unknown variables T_0, T_1, and T_2. Their solutions must be expressions in terms of parameters θ_0, θ_1, and θ_2. Take the first order derivatives on both sides of optimal conditions (A.1), we can obtain the following three equations:

$$\begin{aligned}
B^{''}(T_0+T_1+T_2)\left(\tfrac{\partial T_0}{\partial\theta_0}+\tfrac{\partial T_1}{\partial\theta_0}+\tfrac{\partial T_2}{\partial\theta_0}\right) &= C_0^{''}(T_0;\theta_0)\tfrac{\partial T_0}{\partial\theta_0}+\tfrac{\partial C_0^{'}(T_0;\theta_0)}{\partial\theta_0}\\
B^{''}(T_0+T_1+T_2)\left(\tfrac{\partial T_0}{\partial\theta_0}+\tfrac{\partial T_1}{\partial\theta_0}+\tfrac{\partial T_2}{\partial\theta_0}\right) &= C_1^{''}(T_1;\theta_1)\tfrac{\partial T_1}{\partial\theta_0}\\
B^{''}(T_0+T_1+T_2)\left(\tfrac{\partial T_0}{\partial\theta_0}+\tfrac{\partial T_1}{\partial\theta_0}+\tfrac{\partial T_2}{\partial\theta_0}\right) &= C_2^{''}(T_2;\theta_2)\tfrac{\partial T_2}{\partial\theta_0}
\end{aligned} \tag{A.2}$$

From above three equations in (A.2), we then get the solutions to $\partial T_0/\partial\theta_0$, $\partial T_1/\partial\theta_0$, and $\partial T_2/\partial\theta_0$ as below:

$$\begin{aligned}
\tfrac{\partial T_0}{\partial\theta_0} &= -D_0\,(1-S_0)\\
\tfrac{\partial T_1}{\partial\theta_0} &= D_0 S_1\\
\tfrac{\partial T_2}{\partial\theta_0} &= D_0 S_2
\end{aligned} \tag{A.3}$$

where

$$D_0 = \frac{\partial C_0^{'}(T_0;\theta_0)}{\partial \theta_0} \frac{1}{C_0^{''}(T_0;\theta_0)} \tag{A.4}$$

and for $i = 0, 1, 2,$

$$S_i = \frac{C_i^{''}(T_i;\theta_i)^{-1}}{C_0^{''}(T_0;\theta_0)^{-1} + C_1^{''}(T_1;\theta_1)^{-1} + C_2^{''}(T_2;\theta_2)^{-1} - B^{''}(T_0+T_1+T_2)^{-1}} \tag{A.5}$$

The assumptions that $\partial C_i^{'}(T_i;\theta_i)/\partial\theta_i > 0$, $C_i^{''}(T_i;\theta_i) > 0$, $i = 0, 1, 2$, and $B^{''}(T_s) > 0$ guarantee that $D_0 > 0$, and $0 < S_i < 1$, $i = 0, 1, 2$. Therefore, we have $\partial T_0/\partial\theta_0 < 0$, $\partial T_1/\partial\theta_0 > 0$ and $\partial T_2/\partial\theta_0 > 0$. This implies that, when it becomes more costly to collect tax-0, governments would raise more revenue from tax-1 and tax-2. Assume that taxation is on the "good side" of Laffer curve, more taxation in T_i calls for higher tax rate τ_i.

Moreover, the intensity of the government response in τ_i to the abolition of tax-0 could depend on several factors. Here I discuss three of them relevant to my empirical study: (1) tax sharing ratio of subnational governments; (2) size of tax base; (3) marginal cost of other public funds. Testable predictions will be made for these three cases.

Tax Sharing Ratio

Suppose there are two levels of governments: central and local, and tax enforcement is under the *de facto* control of local governments. Assume tax i is a sharing tax between central and local government, and the sharing ratio for local government is denoted by $\alpha_i \in (0, 1)$. Suppose T_i is the tax revenue *obtained* by local governments. Under tax sharing systems, that means the total tax revenue *collected* for tax-i should be T_i/α. Therefore, the tax collection cost is $C_i(T_i/\alpha_i)$, and we still have $T = \sum_{i=0}^{2} T_i$. Under these conditions, the optimal condition (2.1) now becomes:

$$B^{'}(T) = \frac{1}{\alpha_i} C_i^{'}(T_i/\alpha_i) \qquad ,\forall i = 0, 1, 2 \tag{A.6}$$

Under very general conditions for function form of $C_i(\cdot)$ and $B(\cdot)$, we can prove that the lower the tax sharing ratio α_i, the lower the tax rate τ_i that would be implemented according to Equation (A.1). In addition to that, tax rate τ_i would increase less in response to a given shock from the abolition of tax-0. Geometrically, it is because the $MC_i(\cdot)$ curve not only shifts to the left but also becomes steeper if α_i decreases.

The following Prediction 2 can be made on the effect of tax sharing ratio on tax enforcement under two assumptions.

Assumption 1. $\partial C_0^{''}(T_0;\theta_0)/\partial\theta_0 = 0$. *Any function form like* $C_0(T_0;\theta_0) = f(T_0) + g(\theta_0) + h(\theta_0)T_0$ *would satisfiy this condistion.*

Assumption 2. $B^{'''}(T_s) = 0$, and $C_k^{'''}(T_k;\theta_k) = 0$ for $k = 0, 1, 2$.

Prediction 2. *Suppose that tax-0 is abolished and tax-1's effective tax rate τ_1 increases, under Assumption 1-2, τ_1 would rise by less if the local government has a lower sharing ratio in tax-1.*

Proof: Define $\theta_1 = 1/\alpha_1$. Therefore $C_1'(T_1;\theta_1) = \frac{1}{\alpha_1}C_1'(T_1/\alpha_1) = \theta_1 C(\theta_1 T_1)$. We need to show $\partial^2 T_1/\partial\theta_0\partial\theta_1 < 0$ in order to prove that lower sharing ratio (bigger θ_1) leads to smaller rise in tax-1 following abolition of tax-0.

Further differentiating $\partial T_1/\partial\theta_0$ in (A.2) with respect to θ_1, we obtain

$$\frac{\partial^2 T_1}{\partial\theta_0\partial\theta_1} = \frac{\partial D_0}{\partial T_0}\frac{\partial T_0}{\partial\theta_1}S_1 + D_0\frac{\partial S_1}{\partial\theta_1} \tag{A.7}$$

where $\frac{\partial S_1}{\partial\theta_1}$ satisfies the following expression, which generally holds for $i = 0, 1, 2$, and $j = 0, 1, 2$.

$$\begin{aligned}\frac{\partial S_i}{\partial\theta_j} &= \frac{\partial S_i}{\partial C_j''(T_j;\theta_j)}\frac{\partial C_j''(T_j;\theta_j)}{\partial\theta_j} \\ &+ \sum_{k=0}^{2}\left(\frac{\partial S_i}{\partial C_k''(T_k;\theta_k)}C_k'''(T_k;\theta_k) + \frac{\partial S_i}{\partial B''(T_s)}B'''(T_s)\right)\frac{\partial T_k}{\partial\theta_j}\end{aligned} \tag{A.8}$$

Under Assumption 1, we have $\frac{\partial D_0}{\partial T_0} = 0$. Under Assumption 2, $\frac{\partial S_i}{\partial\theta_j} = \frac{\partial S_i}{\partial C_j''(T_j;\theta_j)}\frac{\partial C_j''(T_j;\theta_j)}{\partial\theta_j}$. Therefore, under Assumption 1-2, Expression (A.7) can be re-written as:

$$\frac{\partial T_1^2}{\partial\theta_0\partial\theta_1} = D_0\frac{\partial S_1}{\partial C_1''(T_1;\theta_1)}\frac{\partial C_1''(T_1;\theta_1)}{\partial\theta_1} < 0 \tag{A.9}$$

In (A.9), $\partial^2 T_1/\partial\theta_0\partial\theta_1 < 0$ because $\partial C_1''(T_1;\theta_1)/\partial\theta_1 > 0$ by assumption, and $D_0 > 0$, $\partial S_1/\partial C_1''(T_1;\theta_1) < 0$ respectively by definition of D_0 in (A.4) and definition of S_1 in (A.5).

Size of Tax Base

Suppose tax revenue $T_i = \tau_i \cdot y_i$ for tax i. Assume the elasticity of tax base to tax rate $\varepsilon_{\tau,i}^y$ is a constant in any value of τ_i for tax i. Then we have the following expression:

$$dT_i/d\tau_i = (1 + \varepsilon_i)y_i \tag{A.10}$$

Assumption 3. *$\varepsilon_{\tau,i}^y$ is a constant and $-1 < \varepsilon_{\tau,i}^y < 0$.*

Assumption 3 implies that the tax is always on the "good side" of the Laffer-curve (that is, $dT_i/d\tau_i > 0$). Equation (A.10) implies that, if the tax base y_i is bigger, then the tax revenue T_i would rise more for the same amount of increase in the tax rate τ_i. Therefore, in order to raise the same amount of tax revenue ΔT_i, a government only needs to raise tax rate by less amount $\Delta\tau_i$ if the tax base y_i is greater.

A prediction can be made as below on the effect of tax base size and the change of tax

rate.

Prediction 3. *Suppose that tax-0 is abolished and tax-1's effective tax rate τ_1 increases, under Assumption 1-3, τ_1 would rise by less if the tax base y_1 is bigger in tax-1.*

Proof: To prove Prediction 3, we need to show $\partial^2\tau_1/\partial\theta_0\partial y_1 < 0$.

Combining Expression (A.4) and (A.10), we have

$$\frac{\partial\tau_1}{\partial\theta_0} = \frac{\partial T_1}{\partial\theta_0}\frac{\partial\tau_1}{\partial T_1} = \frac{D_0 S_1}{(1+\varepsilon_1)y_1} \tag{A.11}$$

Differentiating $\partial\tau_1/\partial\theta_0$ with respect to y_1, we obtain

$$\frac{\partial^2\tau_1}{\partial\theta_0\partial y_1} = -\frac{D_0 S_1}{(1+\varepsilon_1)y_1^2} + \frac{D_0}{(1+\varepsilon_1)y_1}\frac{\partial S_1}{\partial y_1} \tag{A.12}$$

where $\frac{\partial S_1}{\partial y_1} = \left[\frac{\partial S_1}{\partial C_1''(T_1;\theta_1)}C_1'''(T_1;\theta_1) + \frac{\partial S_1}{\partial B''(T_s)}B'''(T_s)\right]\frac{\partial T_1}{\partial y_1}$. Under Assumption 2, we have $\frac{\partial S_1}{\partial y_1} = 0$. Therefore it is straightforward to show that $\partial^2\tau_1/\partial\theta_0\partial y_1 < 0$ by re-writting Expression (A.12) as:

$$\frac{\partial^2\tau_1}{\partial\theta_0\partial y_1} = -\frac{D_0 S_1}{(1+\varepsilon_1)y_1^2} < 0 \tag{A.13}$$

Marginal Cost of Other Public Funds

How does the response of tax rate τ_1 to abolition of tax-0 change if $MC_2(\cdot)$ curve is flatter? In the real world, it implies that the change of marginal cost for public fund-2 must be smaller. A flatter $MC_2(\cdot)$ curve implies a flatter $MC_s(\cdot)$ curve. For a given shock of abolition of tax-0, flatter $MC_s(\cdot)$ results in a lower rise in cost (and also benefit) of public funds in optimality. Consequently, T_1 and τ_1 would increase by less magnitude as a response.

The prediction below summarizes the impact of marginal cost of other public funds on the tax enforcement.

Prediction 4. *Suppose that tax-0 is abolished and tax-1's effective tax rate τ_1 increases, under Assumption 1-2, τ_1 would rise by less if the marginal change of marginal cost of another source of funds (that is, $C_2''(T_2)$) is smaller.*

Proof: To prove Prediction 4, we need to show $\partial^2 T_1/\partial\theta_0\partial\theta_2 > 0$. Similar to the proof of Prediction 2, this can be done by differentiating $\partial T_1/\partial\theta_0 > 0$ with respect to to θ_2 as below:

$$\frac{\partial T_1^2}{\partial\theta_0\partial\theta_2} = \frac{\partial D_0}{\partial T_0}\frac{\partial T_0}{\partial\theta_2}S_1 + D_0\frac{\partial S_1}{\partial\theta_2} \tag{A.14}$$

where $\frac{\partial S_1}{\partial\theta_2}$ satisfies Expression (A.8).

Under Assumption 1, we have $\frac{\partial D_0}{\partial T_0} = 0$. Under Assumption 2, $\frac{\partial S_i}{\partial\theta_j} = \frac{\partial S_i}{\partial C_j''(T_j;\theta_j)}\frac{\partial C_j''(T_j;\theta_j)}{\partial\theta_j}$.

Therefore, under Assumption 1-2, Expression (A.14) can be re-written as:

$$\frac{\partial T_1^2}{\partial \theta_0 \partial \theta_2} = D_0 \frac{\partial S_1}{\partial C_2''(T_2;\theta_2)} \frac{\partial C_2''(T_2;\theta_2)}{\partial \theta_2} > 0 \tag{A.15}$$

In (A.14), $\partial^2 T_1/\partial\theta_0\partial\theta_2 > 0$ because $\partial C_2''(T_2;\theta_2)/\partial\theta_2 > 0$ by assumption, and $D_0 > 0$, $\partial S_1/\partial C_2''(T_1;\theta_1) > 0$ respectively by definition of D_0 in (A.4) and definition of S_1 in (A.5).

B Appendix: VAT Sharing Ratio and the Effective VAT Rate

This section is an appendix to Section 6.1.1. It mainly shows the cross-province variation of the VAT sharing ratio and its persistence over time. Evidence of a relationship between the VAT sharing ratio and the effective VAT rate will also be reported.

Figure B.1 plots the sharing ratio of county government VAT in total province VAT revenue during 2001 - 2007 against 1995 - 2000 data.

The "*VAT share of county government in province*" is defined by "*county government VAT revenue / subnational governments total VAT revenue of a province*".

Similarly, the right panel plots the sharing ratio of provincial governments. The "*VAT share of provincial government in province*" is defined by "*provincial government VAT revenue / subnational governments total VAT revenue of a province*".

I do not plot the sharing ratio of prefectural governments because "*Prefecture Share* = 1 - *County Share* - *Province Share*".

There are two features with Figure B.1. First, the ratios vary considerably across provinces. Second, over time, ratios in most provinces remain stable from 1995 - 2007. That can be seen from the fact that most data points stay around the 45 degree line.

Does the lower sharing ratio of county governments weaken their incentive to enforce VAT and lower the effective VAT rate in that province? Figure B.2 plots the effective VAT rate of each province against "1 - *VAT share of county governments*" (it is also equal to "*Prefecture government share* + *Province government share*"). The effective VAT has been controlled for the (4)-digit industry fixed effects and firm characteristics including export-sales ratio, intermediate input-output ratio, firm size measured by logarithm of value-added, sales, employment, profitability measured by ratio of total profit to sales, firm age, and affiliation, and regional characteristics like GDP per capita and logarithm of GDP.

Figure B.2 shows a slightly but significantly negative relationship between the effective VAT rate and the VAT sharing ratio across provinces.

Table B.1 reports the regression of the effective VAT rate on the "1 - County Share".

Figure B.1: Persistence of VAT Share in Province

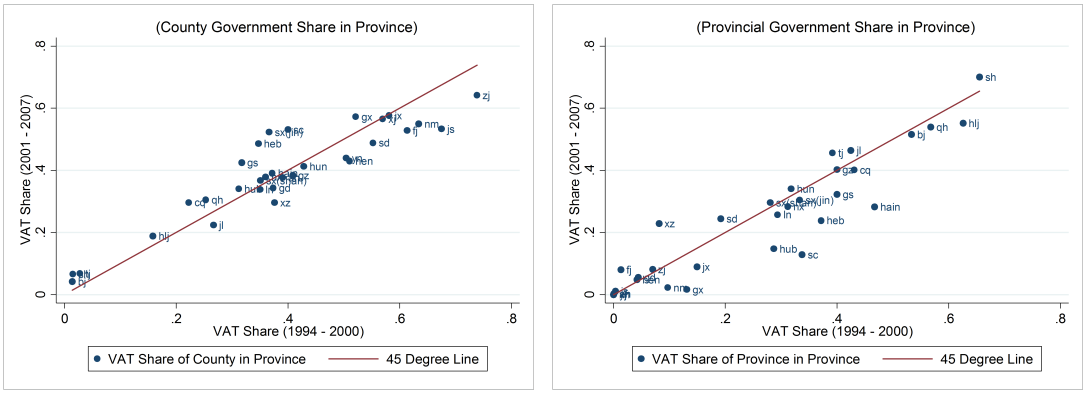
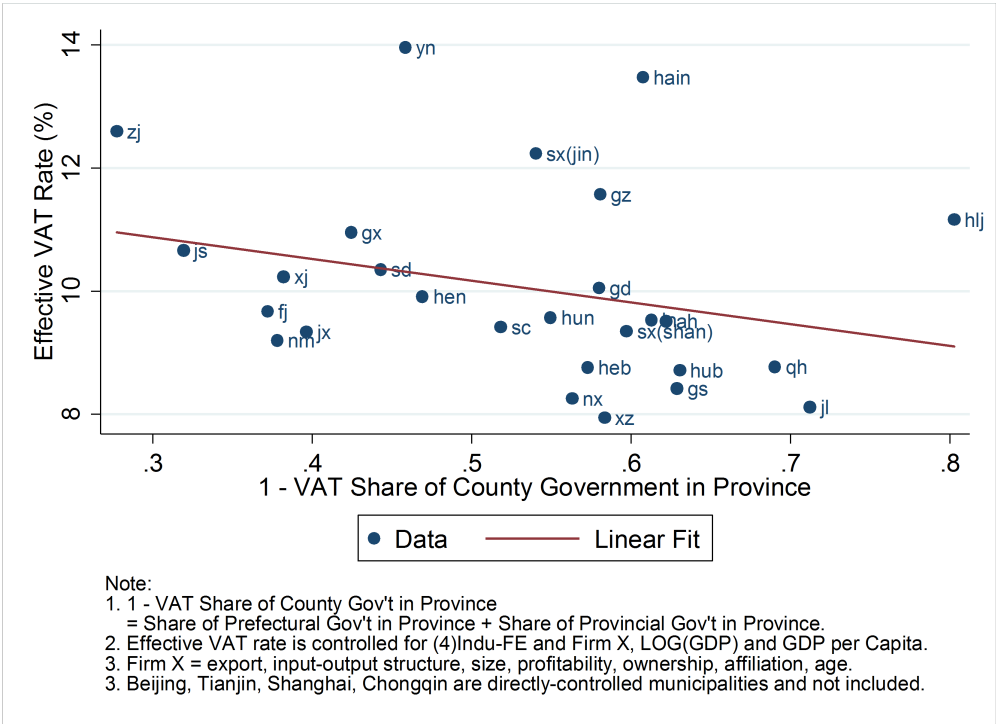


Figure B.2: Effective VAT Rate over VAT Share



Column (1) shows negative coefficient on "1 - County Share", implying the effective VAT rate is higher in provinces where provincial and prefectural governments take more, while county governments take less from 25% of the total VAT revenue belonging to the subnational government. The coefficient value -2.34 implies the effective VAT rate falls by about 0.23 percentage points if the county share rises by 10 percentage points. Clearly, the magnitude of coefficient is so small that a huge part of the effective VAT rate dispersion across regions cannot be explained by the variation in sharing ratio.

To disentangle the impact of province and prefecture share, column (2) and column (3) respectively report the regression on "Prefectural Share" and "Province Share". Both coefficients are negative but not significant at a 10% level. Column (4) includes both "Prefectural Share" and "Province Share", and the coefficients on both are significant at the 5% level.

Table B.1: VAT Sharing Ratio and Incentives for Tax Enforcement Level

	Dependent Variable: Effective VAT Rate			
	(1)	(2)	(3)	(4)
(1 - County Share)	-2.34			
	(0.92)**			
Prefecture Share		-1.37		-3.41
		(1.75)		(1.72)**
Province Share			-1.14	-2.40
			(1.05)	(1.03)**
Controls	*LOG(GDP), LOG(VA), LOG(Sales), Capital/Sales, Ownership, Degree of Competition, Mobility.*			
	Fixed Effects			
Prefecture-Year FE	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Sample size	17,845	17,845	17,845	17,845
R^2	0.12	0.11	0.11	0.12
Clustering	Prefecture level(362 groups)			

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the prefecture level (362 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

C Appendix: Profit Gap Rate

This section is an appendix to the profit gap rate in Section 7.3.2.

Cai and Liu (2009), using the same dataset as this paper, propose using the gap between a firm's imputed profits based on the *National Accounting Principle* (NAP) and its reported accounting profits based on the *General Accepted Accounting Principles*

(GAAP) to measure the degree of tax evasion (or tax avoidance). $Pro_{f,t}$, the reported profit for firm f in year t, can be directly taken from the Annual Survey of Industrial Production. $ImpPro_{f,t}$, the imputed profit for firm f in year t, is calculated in the following way (for notational convenience, I ignore the subscripts):

$$ImpPro = Y - Intermed - FC - Wage - CurDep - VAT \tag{C.1}$$

where Y is the firm's gross output; $Intermed$ is intermediate inputs; FC is financial charges; $Wage$ is the firm's total wage bill; $CurDep$ is current depreciation; and VAT is value-added tax payments.

The profit gap $ProGap = ImpPro - Pro$. To be in line with the measurement of the effective VAT rate. I divide it by value-added to obtain the profit gap rate.

Conceptually, the profit gap can be broken down into four components:

$$\begin{aligned} ProGap = {} & legitimate\ gap + tax\ evasion \\ & + earning\ management + accounting\ error \end{aligned} \tag{C.2}$$

On the right-hand side of expression (C.2), there are three components in addition to *tax evasion*. The first term, *legitimate gap*, is the legal difference between NAP and GAAP.[2] The *accounting error* is assumed to follow a normal distribution with a zero mean. The *earning management* is not a significant concern for the present study as most firms in the sample are not publicly listed and do not have strong incentives to over-report earnings to deceive shareholders.[3] Given Expression (C.2), therefore, we only need to assume that the legitimate gap is not correlated with the tenure of the prefectural secretary of the CPC. Under this assumption, Expression (C.2) allows us to capture some fraction, if not all, of the misreporting of firms' profits.[4]

[2]The GAAP allows for more deductible items in calculating firms' accounting profits than the NAP does. Therefore the accounting profit $Pro_{i,t}$ is generally smaller than the imputed profit $ImpPro_{i,t}$. The *legitimate gap* between the two is approximately equal to: Manufacturing Expenses + Business Taxes and Surcharges + Operating Expense + Management Fees + Asset Impairment Loss + Loss from Changes in Fair Value + Investment Loss + Non-Business Expenditure - Non-Business Income - VAT.

[3]Studies related to this, such as Desai (2003, 2005), are predominantly in corporate finance and accounting literature.

[4]Of course, Expression (C.2) cannot capture tax evasion that simultaneously changes the reported profits and imputed profits by the same magnitude.

Figure C.1: Distribution of Profit-Gap Rate

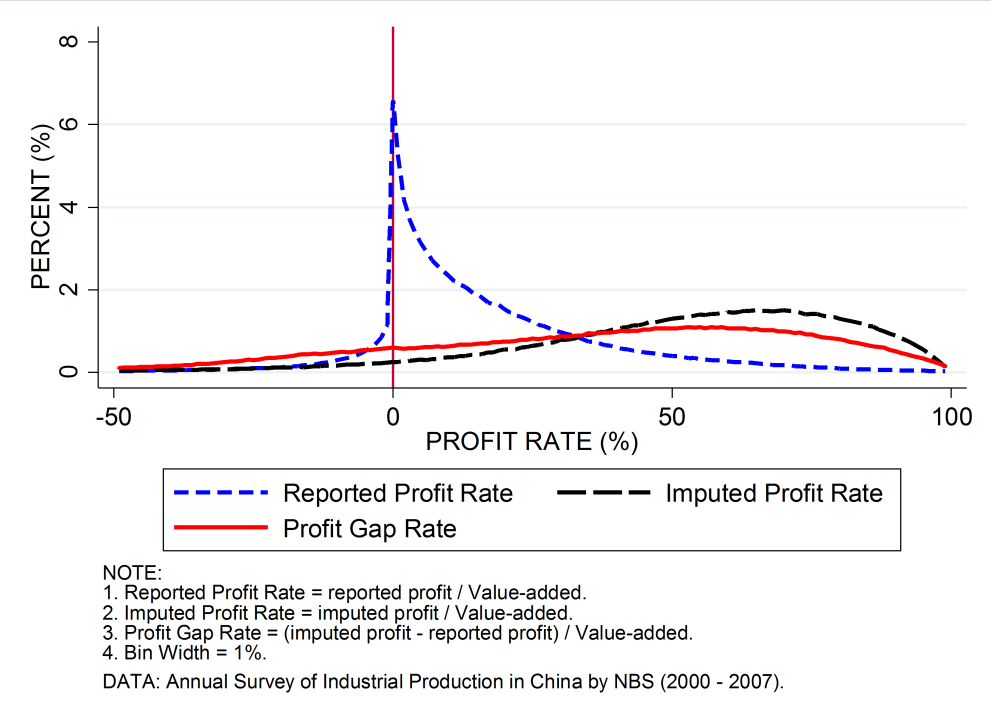


Figure C.1 displays the distribution of imputed profit rate, reported profit rate and the profit gap rate. These three rates are all normalised by firms' value-added. It should be noted that the reported profit rate has a significant spike around zero and is skewed to the right. In contrast, the distribution of the imputed profit rate is quite smooth. The spike suggests that some firms probably report very low but positive profits to evade tax.

Figure C.2: Effective VAT Rate Vs. Profit-Gap Rate

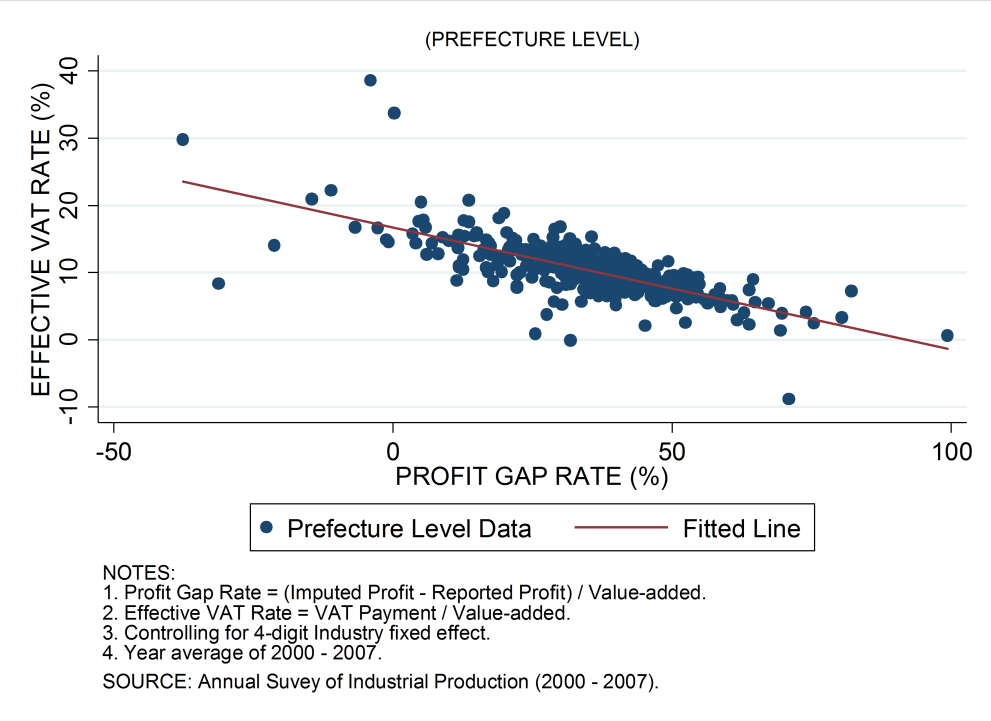


Figure C.2 cross-validates the effective VAT rate and the profit gap rate at the prefecture level after controlling for the 4-digit industry fixed effect. We can see that the two rates are significantly and negatively correlated, implying that tax enforcement may consistently affect both the effective VAT rate and profit misreporting across regions. Since the effective VAT rate and the profit gap rate both have the value-added as their denominators, the negative correlation between the two suggests that the variation of both variables is largely driven by the numerators, which are potentially related to tax evasion, rather than being driven by the denominator.

Figure C.3: Dynamic Effect of AGR. Tax Reform on Profit Gap Rate

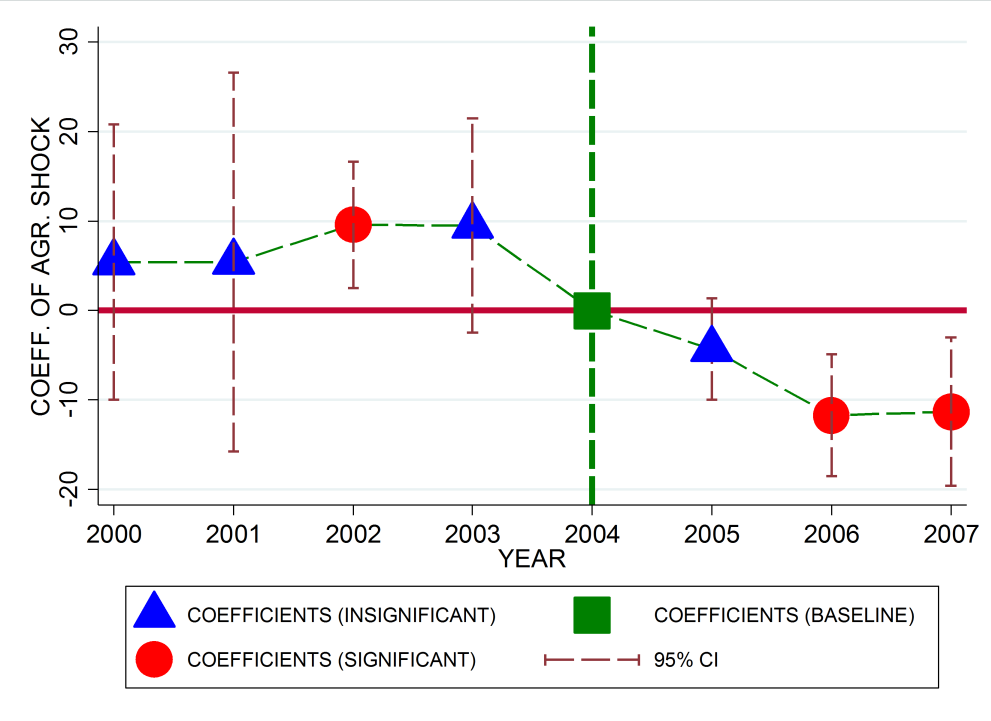
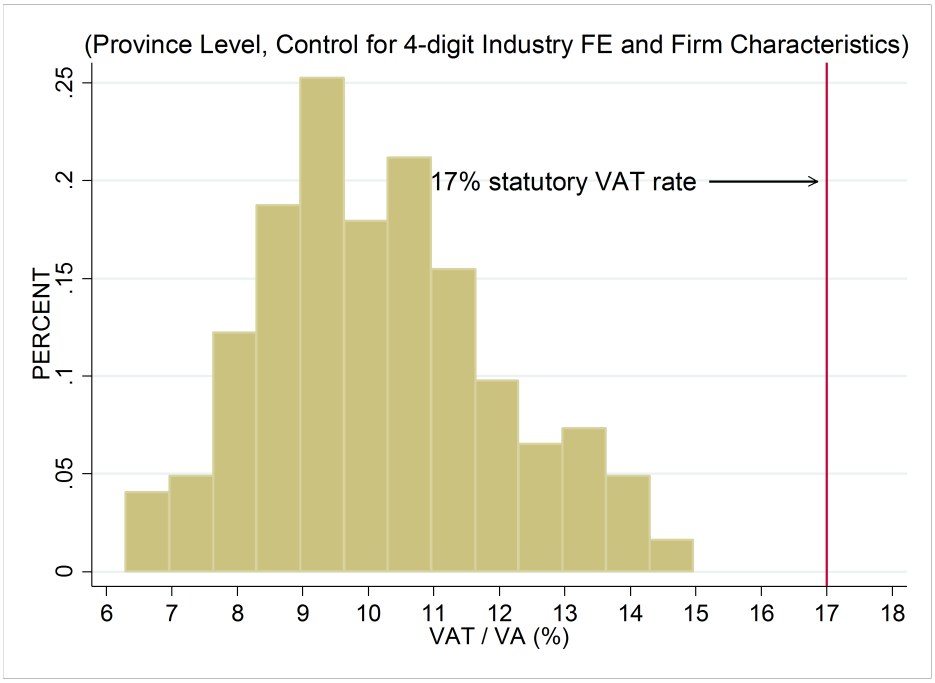


Figure C.3 plots the dynamic effect of the abolition of agricultural tax on the profit gap rate based on Equation (7.2). The effects shown in the figure are consistent with the regression results in Table 8. The fall of the profit gap rate after the agricultural tax abolition suggests enforcements were also strengthened on the corporate income tax. It is harder for firms to evade corporate income tax by under-reporting profits.

D Appendix: Figure

Figure D.1: Distribution of Effective VAT Rate



Notes:
1. This figure controls for two factors that potentially affect the effective VAT rate. First, firm characteristics, including: (1) export/sales, input-output structure measured by intermediate-input/output, ownership type, profitability measured by profit/value-added and Return-on-Asset, firm size measured by log(value-added), log(sales), log(total asset), log(labour), a firm's affiliation type; and, (2) 4-digit industry fixed effect. Second, measurement errors and firm-year specific shocks, which are controlled for by taking the average of the effective VAT rate across firms within provinces and over eight years (2000 - 2007).

2. The dispersion in effective VAT rate across firms could be the result of several factors: (1) Statutory tax codes. Goods may apply at different rates. For example, in China, books and agricultural product are at a 13% low rate. Exports can get the VAT refunded. Some intermediate input, mostly the service, are not deductible. A manufacturing firm F may outsource the logistics service to another firm S in the service sector. In this case, firm F's effective VAT rate = (sales of firm F $\times$ 17 % - 0) (sales sales of firm F - purchased from firm S) $>$ 17%. (2) measurement error and yearly idiosyncratic shock to firm. For example, a firm may buy a huge quantity of coal as intermediate input this year, therefore it would stir the VAT payment over time and make the effective VAT rate low in one year and high in the next. (3) tax enforcement and tax evasion.

Figure D.2: Distribution of Effective VAT Rate (Raw Data)

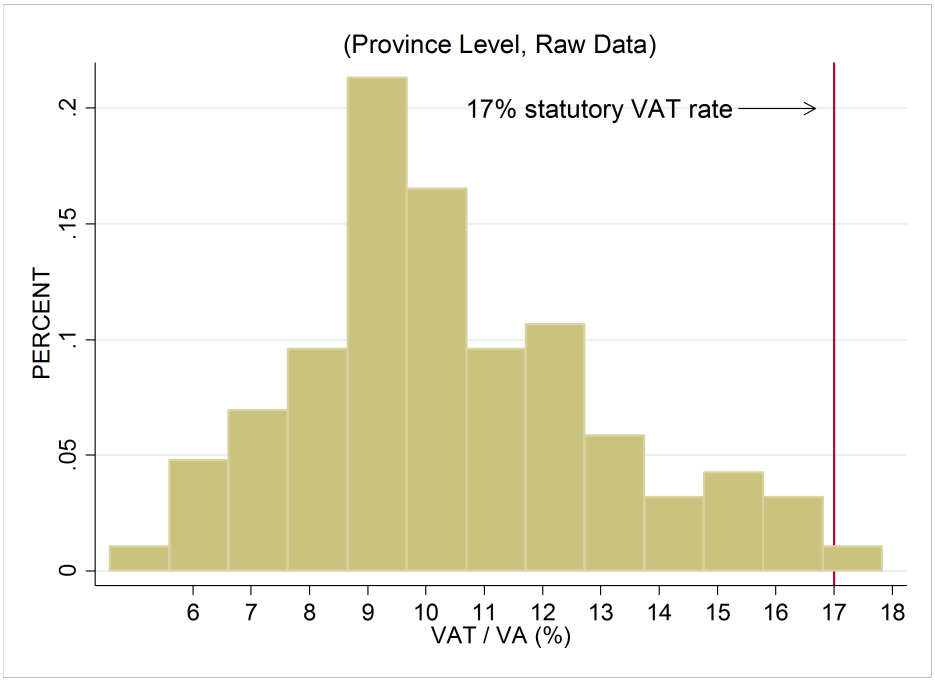


Figure D.3: Distribution of Effective VAT Rate (Firm Level, Raw Data)

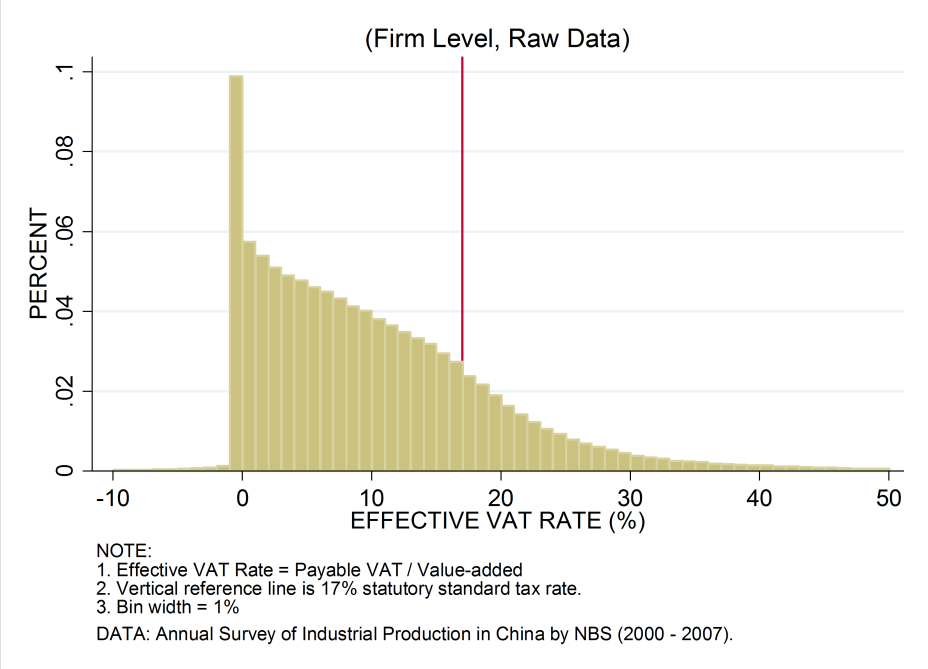


Figure D.4: Underground Sales of Fake Invoices

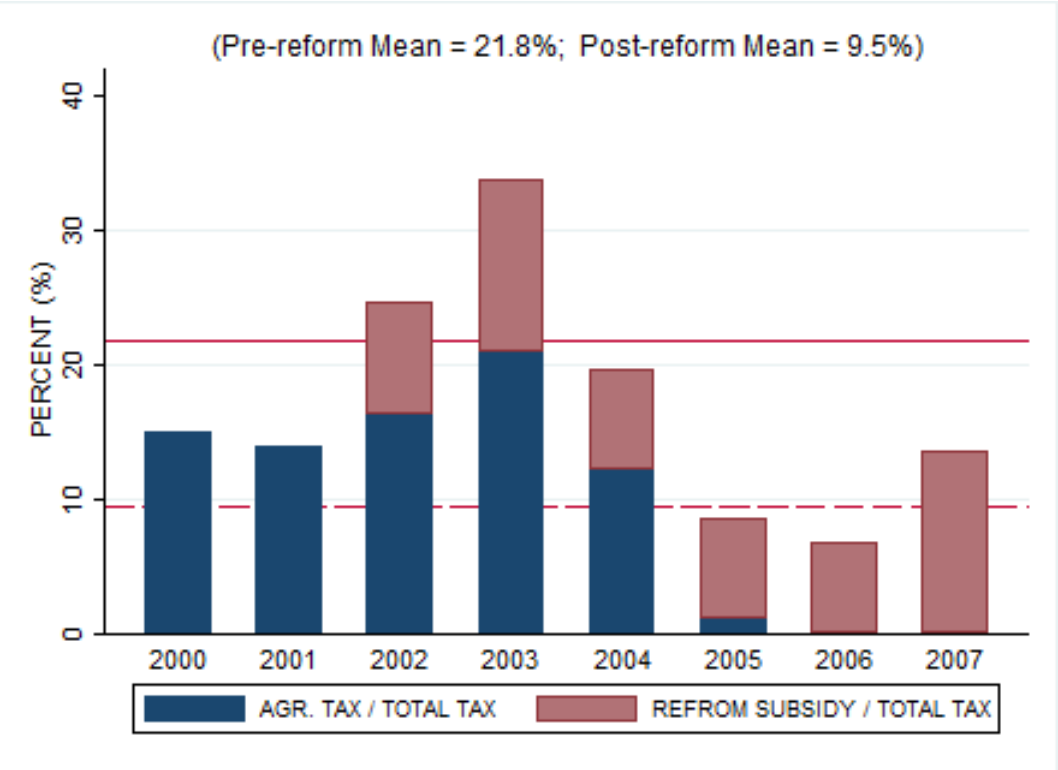
Notes:

1. Both the left and right panel are a print-screen from the author's mobile phone. They show the conversation via text message between the author and two fake invoice dealers in China.

2. The left panel shows a text message between the author and a fake invoice dealer in Kunming, Yunnan Province. The message is a spam advert sent out by the fake invoice deader thirty minutes after the author arrived at the airport in Kunming, April 26, 2014. The following is a translation to English. (Dealer) Hello! Our company can issue various kinds of (invoices from, note: this are the implied words form the context) the State Administration of Taxation, Local Administration of Taxation, for the service sector, engineering, accommodation. Checkable on-line! Good price! Tel: 13078282477. Contact: HuiLin Yang. (Author) Can you issue VAT Special Invoices for me? Can your invoice number match with that of the Administration of Taxation? Where are your invoices from? From Administration of Taxation or from other sources? (Dealer) We have Guanxi (social connection) and do not pay tax. So it is cheap.

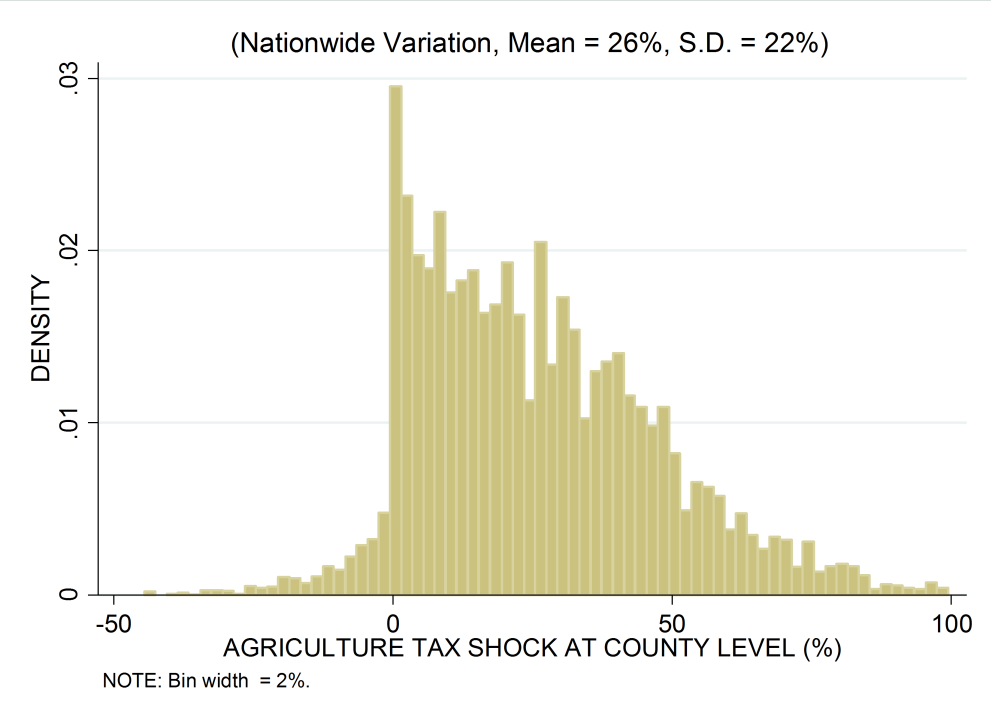
3. The right panel shows the text message between the author and another fake invoice dealer in Beijing, the capital city of China. The message is a spam advert received when the author was in Beijing in April, 2014. The head of the screen shows the phone number of the invoice dealer. Below is the translation of part of the conversation. (Author) Can you issue legal VAT special invoices? How much point do you charge? (Dealer) Eight points (Note: that means the dealer charges 8% of the face value of an invoice issued). (Author) Is it legal invoice? Can your invoice number match with that of the Administration of Taxation? Is there any limitation on the types of goods being purchased and sellers of the goods that are recorded on invoices? Thanks! (Dealer) No limitation, whatever you want. (Author) I am referring to the VAT special invoices. Is there still no limitation on the type of goods and the sources of product origin so that you can issue the invoices as to whatever I demand? (Dealer) Yes.

Figure D.5: Agricultural Taxes and Government Transfer over Time



Notes:
1. This figure displays the county level over-time variation in the agricultural tax revenue and subsidy associated with agricultural tax reform as ratios of total tax revenue.

Figure D.6: Distribution of Agricultural Revenue Loss across Counties



Notes:
1. This figure plots the distribution of revenue loss due to abolition of agricultural tax across almost all county-level jurisdictions in China. The revenue loss is calculated based on Expression (4.3).
2. By 2007, there were 2,860 county-level jurisdictions according to National Bureau of Statistics. In this figure, there are 2,917 county-level jurisdictions. This is because of the Special Economic Development Zone or Science and Technology Park in some places which may be recognised by local governments as special county-level jurisdictions and are assigned a 6-digit county region code.
3. The revenue loss to a small number of counties are negative, showing that these counties could have benefitted from the reform as they received more in subsidy than they lost in taxes following the reform.

Figure D.7: Geographical Distribution of Agricultural Revenue Loss across China

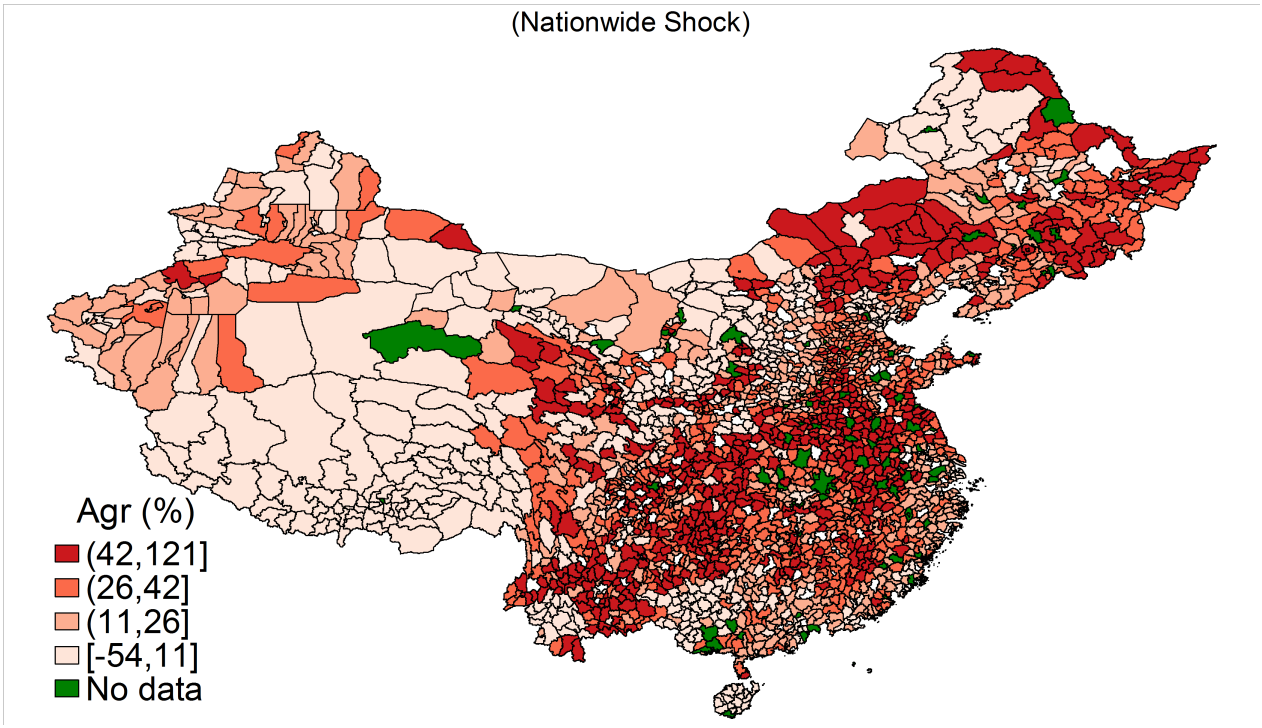


Figure D.8: Distribution of Agricultural Revenue Loss within Prefecture

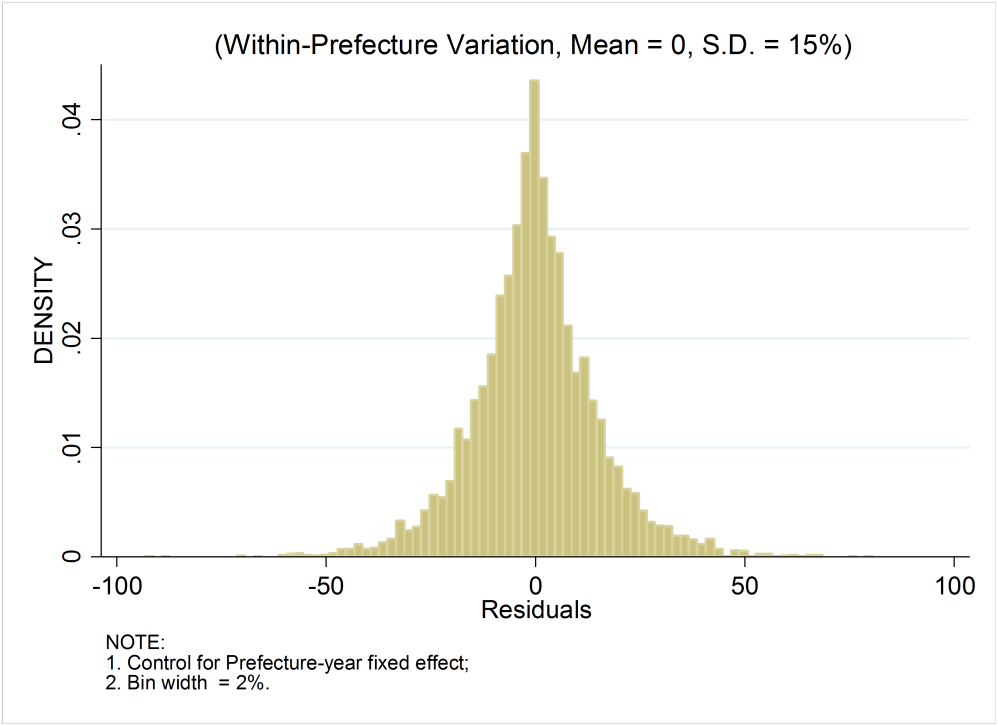


Figure D.9: Agricultural Revenue Loss within Aba Prefecture in Sichuan

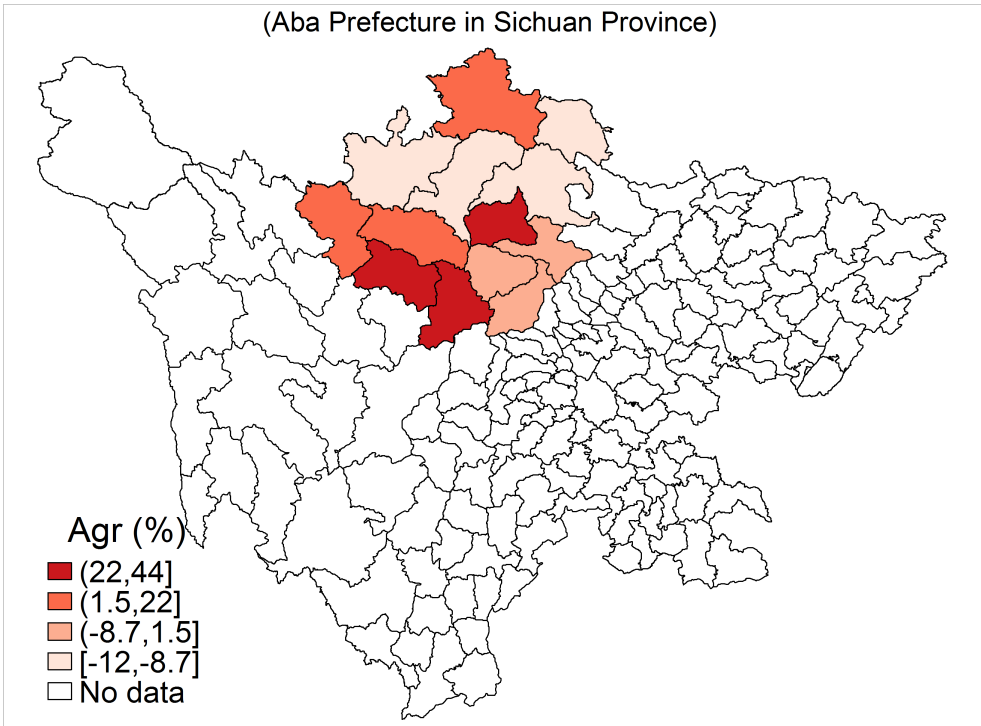


Figure D.10: Effective VAT Rate over Time (Grouping by Median)

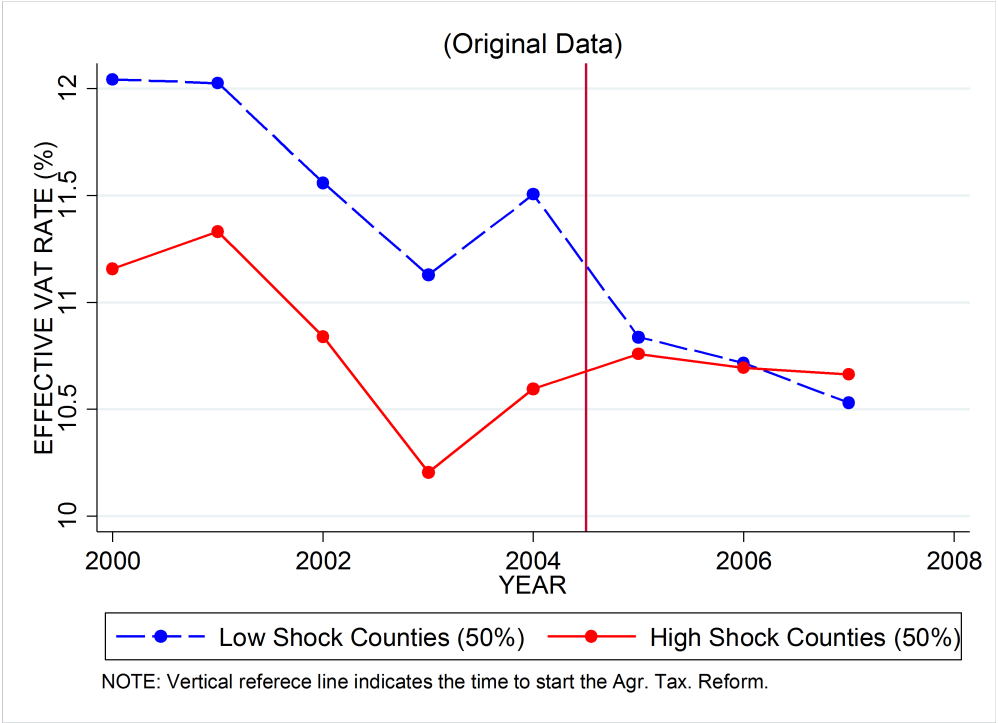


Figure D.11: Effective VAT Rate over Time (Grouping by 1st Quartile)

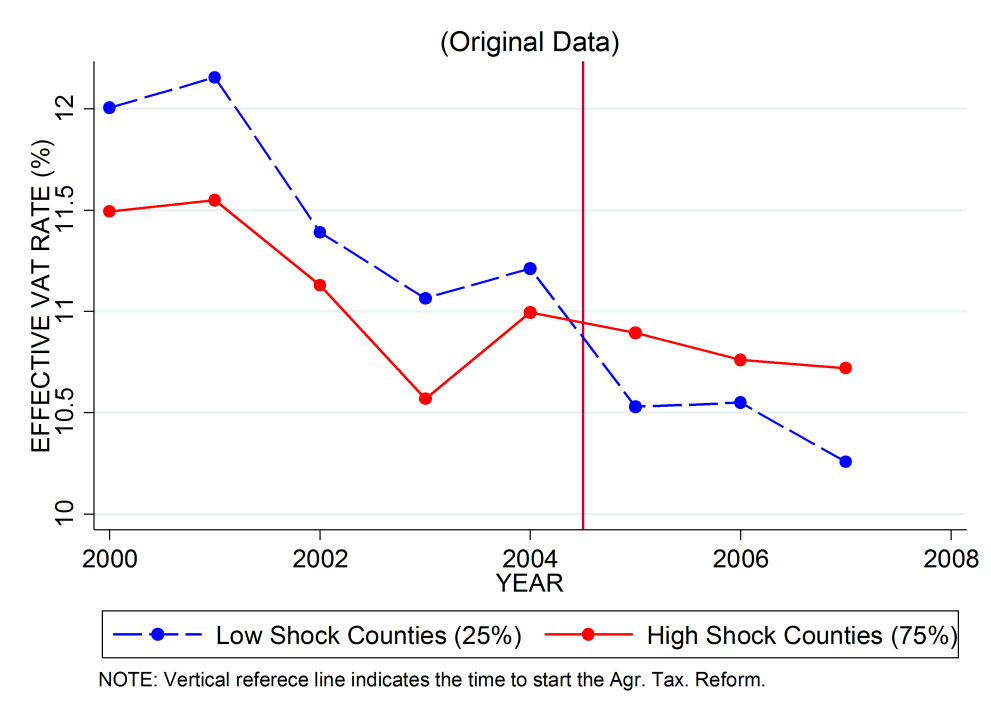


Figure D.12: Effective VAT Rate over Time (Grouping by 1st Quartile)

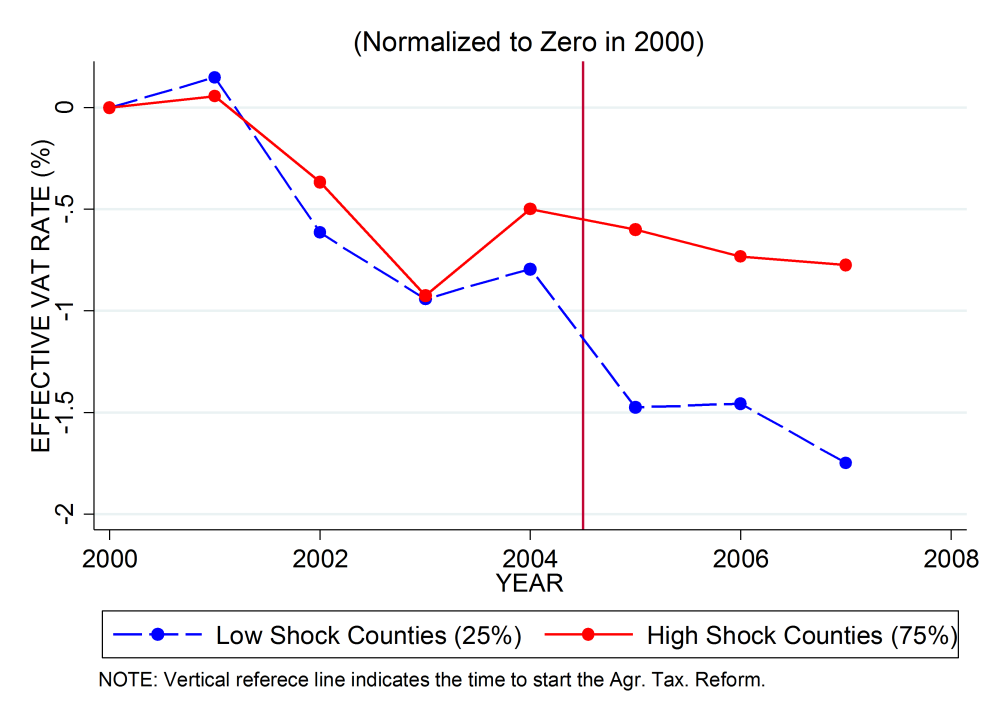


Figure D.13: Distribution of Revenue Loss in Pilot Province – Anhui

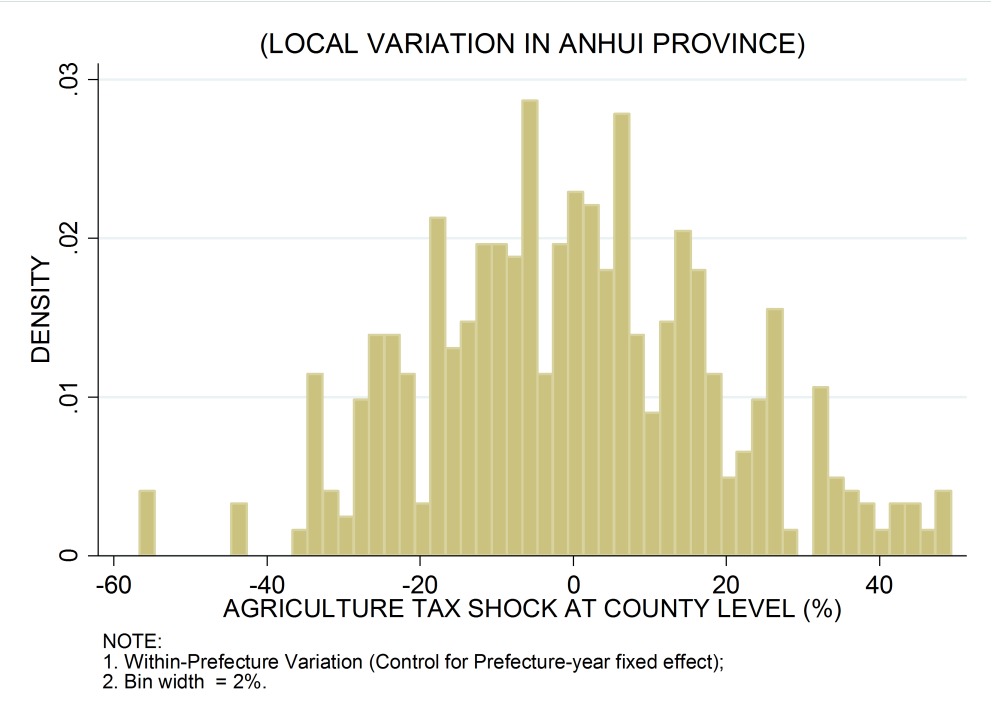


Figure D.14: Revenue Loss in Pilot Province – Anhui

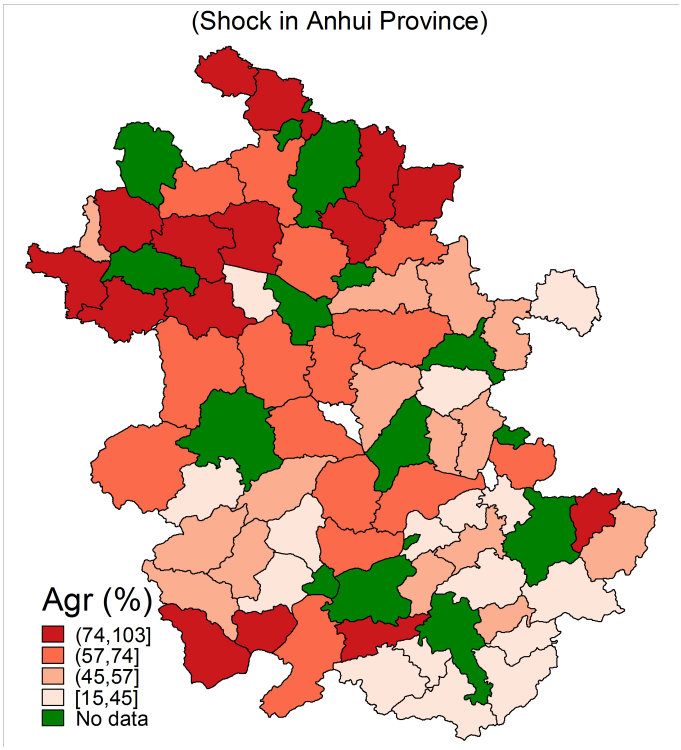


Figure D.15: Distribution of Revenue Loss in Pilot Province – Jiangsu

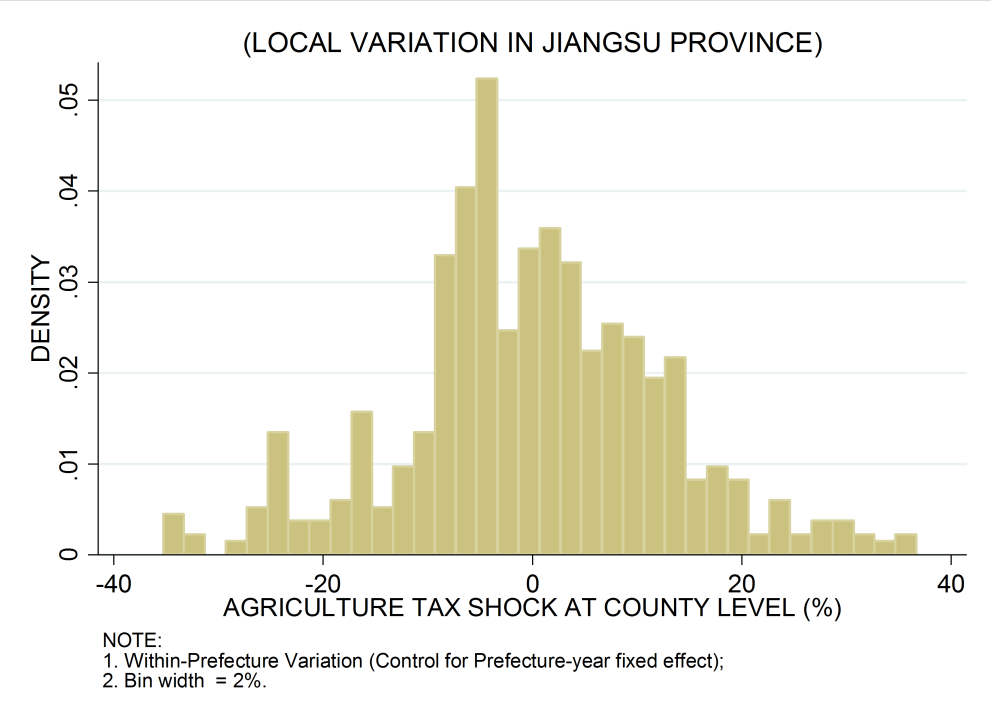
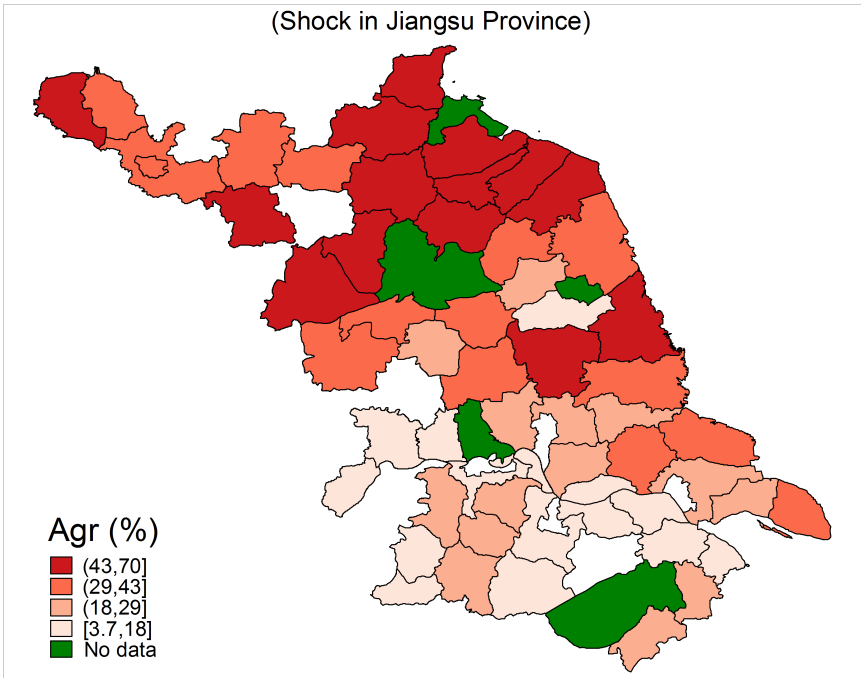


Figure D.16: Revenue Loss in Pilot Province – Jiangsu



Chapter 2

Political Determinants of Selective Tax Enforcement: Evidence from China*

Abstract

Politicians can play a critical role in administration of tax and other policies, but this role is rarely documented or investigated in quantitative approaches. In this paper I first study VAT administration in China, which is the *de jure* responsibility of the State Administration of Taxation but is subject to the *de facto* influence of local politicians, particularly the prefectural secretaries of the Communist Party. Using the variation in turnover of secretaries between 2000 and 2007, I find that, over the tenure of the prefectural secretaries, the effective VAT rate changes in favour of capital-intensive industries, and to the detriment of labour-intensive industries. Additional evidence reveals that the favouritism towards capital intensity is not limited to VAT enforcement, but is also present for corporate income tax and access to credit. I conclude the paper by discussing several possible channels for this favouritism over the tenure of prefectural secretaries. The evidence seems to be most consistent with the explanation of corruption.

*The author is grateful to Francesco Caselli and Camille Landais for their supervision throughout the research on this project and other related topics. Acknowledgements go to Tim Besley, Torsten Persson, Ethan Ilzetzki, Steve Pischke, Henrik Kleven, Michael Peters, Johannes Spinnewijn, Robin Burgess, John Sutton, Frank Cowell for their suggestions and comments, and to Eddy Tam for very helpful discussions. The author also appreciates the feedback from participants in Work-in-progress Seminars on Public Economics, Macroeconomics, and Labour at LSE, and in Seminars at Chinese Academy of Social Sciences, Renmin University of China, Central University of Finance and Economics. Special thanks to Songtao Tan for sharing the dataset on Turnover of Prefectural Secretary, and Heng Yin for data on County Public Finance Statistics. The author, of course, bears the responsibility for all the errors in the paper.

1 Introduction

De jure rates of taxation on corporations and individuals are predictable, codified, and based on a number of observables. However, tax evasion implies that *de facto* tax rates and their distribution across firms are quite different from *de jure* tax rates and their distribution. But what explains the distribution of *de facto* tax rates across economic agents? And how do political institutions and corruption affect the incentives for selective enforcement by tax authorities? Does favouritism towards corruption-susceptible firms inevitably lead to harsh policy at incorrupt firms? And is this pattern of selective enforcement common as well to other maneuverable policies such as credit allocation?

In this paper I begin my study with the distribution of *de facto* VAT rates across industries in China. The enforcement of VAT is the responsibility of local offices of the State Administration of Taxation (SAT), but is subject to the influence of local politicians, particularly the prefectural secretaries of the Communist Party. I find that in the early years of a secretary's tenure in a certain locality, *de facto* VAT rates are relatively high for local firms in capital-intensive industries. However, the longer the same individual retains the role of local party secretary, the more the *de facto* tax burden shifts to labour-intensive firms. Quantitatively, additional year of prefectural secretaries' tenure is associated with a drop of the effective VAT rate by 8% for capital-intensive firms, and a rise by 3% for labour-intensive counterparts.[1]

In a similar vein, I use a natural experiment that creates variation across counties through a "fiscal squeeze" that forces them to tighten up enforcement in order to bolster revenues, to show that the longer the current prefectural secretary has been in his/her role, the more the burden of adjustment falls on labour-intensive rather than capital-intensive firms.

I generalize my findings by showing that two additional firm-level outcomes which are potentially susceptible to political manipulation are sensitive to the local secretary's tenure. First, the profit gap – a proxy for the misreporting of company profits – widens by 25% (shrinks by 2.3%) for capital- (labour-) intensive firms as the party secretary's tenure increases. This suggests that monitoring becomes more lax for capital-intensive firms, and also indicates that the tenure effect may extend to evasion of forms of tax other than VAT. Second, I find that over a secretary's tenure, capital- (labour-) intensive firms are able to accumulate more (less) debt. Since the allocation of credit in the economy in China is subject to considerable political influence, this suggests that increasing favouritism towards capital-intensive firms over the tenure of a secretary is not limited to less aggressive tax enforcement.

[1] In this paper, capital-intensity of an industry, denoted by $CapInt$, is measured by the median of firms' capital-sales ratio of the industry. After normalization, an industry with capital-intensity equal to one (zero) is called a capital- (labour-) intensive industry. Readers can refer to Table 1 for the measurement of $CapInt$.

I conclude the paper with a discussion of possible explanations for the tenure effect. I argue that the most plausible explanation is corruption. In support of this explanation I present some circumstantial evidence. Following Cai et al. (2011), I use the ratio of a company's entertainment and travel costs (ETC) to its sales as an indicator of corruption. I find that capital-intensive firms have increasingly higher ETC relative to sales over the tenure of a prefectural secretary. In contrast, the ETC-to-sales ratio of labour-intensive firms remains roughly constant. Further evidence shows that capital-intensive industries are incentivized to participate in corruption for lucrative returns. For each RMB yuan of expenditure on ETC, capital-intensive firms can obtain both greater VAT reductions and easier access to finance than their labour-intensive counterparts. Quantitatively, for capital-intensive firms, ten percentage points increase in the ETC-to-sales ratio is associated with a reduction by 0.2 percentage point in the effective VAT rate and a drop by 0.71 percentage point in the probability of collateral requirement for bank loans. In contrast, the return of ETC for labour-intensive firms is not robustly significant.

However, I cannot entirely rule out alternative explanations, such as politicians' preferences, industrial policies, political connections and learning effects. I take these possibilities into consideration in my empirical work. Nevertheless, it seems difficult to fully rationalize the evidence with these explanations.

Figure 1.1: Distribution of Effective VAT Rate

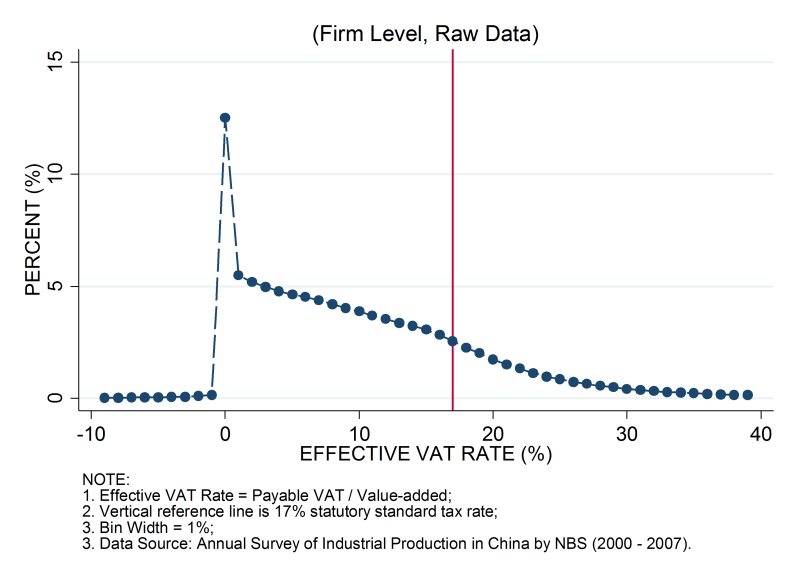


Figure 1.1 plots the distribution of the effective VAT rate across firms in a histogram. It shows that the dispersion of the actual VAT rate is huge, even though the statutory standard tax rate is 17% for the whole country.[2] Of course, this dispersion in the effective VAT rate across firms may be due to legitimate reasons such as lower rates for agricultural

[2]It also should be noted that the effective VAT rate for some firms may exceed 17%. This is because intermediate inputs may not be deductable for these firms.

products and export refunds.[3] However, the huge variation in the effective VAT rate can not be fully explained by legal factors, even after controlling for a long list of firm characteristics which could potentially affect the effective tax rate.[4]

To further pin down the incentives for local political intervention, this paper attempts to study the cross-industry distribution of the effective VAT rate over local political cycles. It tries to show whether and how the cross-industry distribution of VAT changes over the tenure of a prefectural secretary of the Communist Party.

Relations to the Literature

The paper finds rising favouritism towards capital-intensive industries and simultaneous harsher treatment on labour-intensive ones as the prefectural secretary of the Communist Party stays in office longer. The finding is complementary to a number of related studies on corruption and political connection by drawing a bigger picture of their impact not only on the corrupt or connected firms, but also on the incorrupt and unconnected ones. In contrast, most existing studies have only shown the direct benefits brought about to corrupt or connected firms. For example, political connections affect the taxation of firms, in China and other countries (Adhikari et al., 2006; Faccio, 2007; Wu et al., 2009). In addition to tax, there are other preferential policies – such as credit, permission to access profitable markets and so on – that corrupt or politically connected firms can obtain from government officials (Fisman, 2001; Fisman and Wang, 2015, 2014; Faccio, 2006, 2007; Khwaja and Mian, 2005, 2011; Fan et al., 2007; Li et al., 2008). Some papers have pointed out that discriminative bank loans are prevalent in China (Li et al., 2008; Allen et al., 2005; Ponet et al., 2010; Cull and Xu, 2003; Fith et al., 2009). Moreover, the tenure effect identified in this paper is also consistent with the finding of Koren et al. (2014) that politically connected firms in Hungary win more procurement contracts, but that the number of contracts only gradually increases after the associated party wins the election, rather than experiencing an immediate jump.

The paper may also shed some light on the distortion and misallocation issues in China. The results show that the tenure effect of a prefectural secretary on the VAT rate actually reduces the dispersion of the tax rate between capital-intensive and labour-intensive industries, which should be regarded as a good signal for an improvement in aggregate productivity a la Hsieh and Klenow (2009). Of course, this may not be the case

[3]Here I introduce a simple case as an illustration. A manufacturing firm F may outsource the logistics service to another firm S in the service sector. Acoording to the tax law in China, firm S in the service sector is paying business tax and cannot issue the VAT special invoices to firm F. In this case, firm F's effective VAT rate = (sales of firm F $\times$ 17 % - 0) / (sales sales of firm F - purchased from firm S) $>$ 17%.

[4]These characteristics include: exports/sales, input-output structure measured by intermediate-input/output, ownership type, profitability measured by profit/value-added and return-on-assets, firm size measured by log(value-added), log(sales), log(total asset), log(labour), and firm's affiliation type.

if we also include distortions other than VAT. Further studies need to be carried out in this regard.

Based on cross-country data, some recent literature stress the complementarity between legal capacity and fiscal capacity (Besley and Persson, 2009, 2011), as well as the interaction between economic development, tax system, and political institutions (Besley and Persson, 2013). Additionally, most studies on tax enforcement in developing countries have long emphasized the information problem and the role of the informal sector in undermining taxation capacity (Gordon and Li, 2009). This paper provides micro evidence on how the weak legal capacity of a nation could translate into a weak taxation capacity through the channel of political intervention even where information and informality are not a problem. The paper suggests that a good taxation technology like VAT could be crippled by unsatisfactory political institutions, and impeccable law codes might be distorted by corruption.

Around the world, VAT has long been thought of as one of the best taxation "technologies" in use to date. For example, Pomeranz (2013) provides evidence from a field experiment on the advantage of VAT in revealing information along the value-added chain. The findings in this paper, however, demonstrate that VAT in China is not operating at its "technological frontier". This finding is consistent with some cross-country studies which find that VAT administration is not efficient in developing countries. For example, Keen and Lockwood (2010) find that the revenue impact of the introduction of VAT is actually negative for poor countries. Aizenman and Yothin (2008) find that the political instability of a country reduces its VAT collection efficiency. Chen (2015a) suggests that the within-industry dispersion in the effective VAT rate leads to a loss in China's manufacturing sector of about 7% of TFP. Chen (2015b) shows that VAT enforcement in China can be affected by shocks on local governments budget.

There is a huge literature on tax evasion. Andreoni et al. (1998), Slemrod and Yitzhaki (2002) and Slemrod (2007) provide comprehensive reviews in this regard. Slemrod et al. (2001) and Kleven et al. (2011) use field experiments to study tax-payers' behavioural responses to the threat of auditing. Tax evasion is not a trivial issue in China. Fisman and Wei (2004) find that importing firms in China may evade tariffs and VAT by mislabelling the classification of imported products from Hong Kong. Cai and Liu (2009) find that market competition may make firms more likely to conduct tax evasion by underreporting their profits. Alumnia and Lopez-Rodriguez (2014) highlight the cross-industry heterogeneity of firms' misreporting in response to tax enforcement. Most of the existing literature highlights the incentive of tax-payers in tax evasion and emphasises the information problem. Khwaja et al. (2014) is one of the few papers that studies how pecuniary incentives could stimulate tax inspectors' efforts. My paper focuses on the role of politicians in tax administration and its impact on the effective tax rate.

Paper Structure

The rest of the paper is organized as follows. Section 2 sets up a model of optimal tax enforcement across industries where corruption is available. Section 3 introduces the institutional background. Section 4 describes the datasets and measurements of variables. Section 5 presents the main empirical strategies and the results of the tenure effect on the effective VAT rate. Section 6 studies the tenure effects on two other tenure-sensitive outcomes: profit misreporting and the debt ratio. Section 7 discusses the possible underlying mechanisms of the tenure effect on taxation and credit allocation. Section 8 concludes.

2 Model – Optimal Tax Enforcement with Corruption

This paper focuses on the capital-intensity of industries because it is widely suggested that, although corruption might be pervasive around the world, it is quite concentrated in certain industries. By and large, more corrupt industries turn out to be more capital-intensive than less corrupt ones.[5] The relationship between an industry's capital-intensity and its corruption intensity gives rise to two following questions. First, why are the capital-intensive firms more susceptible to corruption than the labour-intensive ones? Second, if some corrupt firms can obtain a tax reduction or receive other preferential treatment, does it necessarily mean that other incorrupt ones have to pay more tax and suffer harsh treatment against them? The model in this section is aimed to answer these two questions and demonstrate underlying mechanisms.

The model has two features. First, the government has to face a budget constraint. Second, there exists homogeneous fixed cost of corruption and heterogeneous return from corruption across industries. The return includes tax reduction and others such as easy credit and cheap land. In the paper, the model is kept as simple as possible for illustrative purposes. I will not comprehensively analyse and discuss the model. Propositions of the model are presented only for the cases relevant to the empirical findings in the paper. It should be noted that although the model focuses only on the tax policy, it can easily extend to other government policies, such as allocation of credit and land, where resource constraints are binding for the government.

The answer to the second and first question are summarized by Proposition 2 and Proposition 3 respectively. Proposition 2 shows that the distribution of the effective tax rates across industries may change as the overall business environment grows more

[5] It is shown that, among other all sectors, the following manufacturing sectors are more prone to graft: (1) coal, palm oil and timber; (2) oil, gas, chemicals and other energy; (3)steel, other metals, mining and commodities (Economist, 2014). Most all of them are capital-intensive manufacturing sectors. Similarly, according to a study based on 427 cases of bribery in international business by the OECD, two-thirds of the cases occurred in just four industries: extractive (19%); construction (15%); transportation and storage (15%); and information and communication (10%) (OECD, 2014).

corrupt. Specifically, the tax rate drops for the corruption-susceptible capital-intensive industries and rises for the labour-intensive ones. This is mainly due to the assumption that government has to face a budget constraint. Proposition 3 suggests that the corruption of capital-intensive industries are more responsive to tax rate because they can get more other benefits, such as bank credit, from corruption in addition to tax reduction. The underlying mechanism is, given the same fixed cost of corruption, bigger other benefits from corruption can complement the benefit of tax reduction and the complementarity strengthens incentives of capital-intensive firms to participate in corruption.

Empirical tests of the models require exogeneous variation in the overall business environment and rent-seeking opportunity. In most parts of the paper following Section 2, I will argue, with both reasoning and empirical facts, that the turnover and tenure of a prefectural secretary of the Communist Party is potentially a good source of such variation.

2.1 Optimal Effective Tax Rates across Sectors

In this part I characterize the optimal effective tax rates across sectors in the presence of tax collection costs and corruption. The general equilibrium effect of corruption on taxation can be rationalized by the optimal conditions for taxation.

There are two types of agents in this model: tax-payers (firms) and the tax administrator. There are N industries, each indexed by i, and one unit continuum of ex ante homogeneous firms in each industry, with each firm indexed by j.

Tax administrators can choose to enforce the effective tax rate τ_i for each industry with costly auditing. For any given τ_i, suppose there are a fraction $1-u_i(\tau_i)$ of firms in industry i that choose to bribe government officials (not necessarily the tax administrators). In return, the bribing firms only need to pay a tax rate $\bar{\tau}_i$, which is lower than τ_i. For simplicity, $\bar{\tau}_i$ is exogeneously given. Obviously, the higher the tax rate, the more incentive to bribe the firm has. Therefore, $u_i^{'}(\tau_i) < 0$.

In choosing the effective tax rate τ_i, the tax administrator faces a trade-off between the benefit of public funds and taxation costs, as well as his/her own personal gain. The trade-off can be summarized by the objective function as below:

$$max_{\{\tau_i\}_{i=1}^N} \qquad B(T) - \sum_{i=1}^{N} \tilde{C}_i(\tau_i) + \sum_{i=1}^{N} V(w)$$

where $B(T)$ is the total benefit of public funds T, with $B'(T) \leq 0$, $B''(T) \geq 0$. $T = \sum_{i=1}^{N} \tilde{T}_i$, where $\tilde{T}_i$ is the total tax payment in industry i. $\tilde{T}_i = u(\tau_i)T_i + (1-u_i(\tau_i))\bar{T}_i$, where T_i and $\bar{T}_i$ are the tax revenue of each non-bribing firm and bribing firm in industry i, respectively, with $T_i = \tau_i Y_i(\tau_i)$ and $\bar{T}_i = \bar{\tau}_i Y_i(\bar{\tau}_i)$. $Y_i(\tau_i)$ reflects the relationship between the tax base and the tax rate. It is a reduced form that includes all types of response of

a firm's tax base Y_i to tax rate τ_i, with $Y_i'(\tau_i) \geq 0$. Suppose the tax rate is always on the good side of the Laffer Curve (augmented with corruption), that is, $\tilde{T_i'}(\tau_i) > 0$.

$\tilde{C}_i(\tau_i)$ is the tax collection cost in industry i when the tax rate is τ_i. It can be expressed as below:

$$\tilde{C}_i(\tau_i) = u_i(\tau_i)C_i(T_i) + (1 - u_i(\tau_i))\, C_i(\bar{T}_i)$$

where $u_i(\tau_i)C_i(T_i)$ is the tax collection cost for non-bribing firms, and $(1 - u_i(\tau_i))\, C_i(\bar{T}_i)$ for bribing ones. The tax collection cost function $C_i(T_i)$ satisfies $C_i^{'}(T_i) > 0$ and $C_i^{''}(T_i) > 0$.

$V(w)$ is the tax administrator's utility obtained from income w, and w can be written in the following way:

$$w = w_0 + \theta\,(1 - \beta) \sum_{i=1}^{N} (1 - u_i(\tau_i)) \left(T_i - \bar{T}_i\right)$$

where w_0 is the non-corruption income. $\theta\,(1 - \beta)\,(1 - u_i) \left(T_i - \bar{T}_i\right)$ is the rent that the tax administrator obtained from the bribing firms. $(1 - u_i) \left(T_i - \bar{T}_i\right)$ is the total rent of corruption in industry i. Parameter $\beta \in (0, 1)$ reflects the Nash bargaining power of bribing firms in rent-seeking. $(1 - \beta)\,(1 - u_i) \left(T_i - \bar{T}_i\right)$ is the total bribes paid by firms. Parameter $\theta \in [0, 1]$ is the share of bribes that the tax administrator can really obtain. θ might be less than 1 because the corruption is conducted through a political network. Everyone involved eventually take only a slice from the bribes.

There are three elasticities that are useful to describe the optimal conditions. They are defined as below.

Definition. (1) The elasticity of tax collection costs to the tax rate: $\epsilon^{c}_{\tau,i} = \frac{C_i^{'}(T_i)}{C_i(T_i) - C_i(\bar{T}_i)} \left(T_i - \bar{T}_i\right)$; (2) the elasticity of a firm's tax payment to the tax rate: $\epsilon^{T}_{\tau,i} = \frac{T_i^{'}(\tau_i)}{T_i(\tau_i) - \bar{T}_i} (\tau_i - \bar{\tau}_i)$; (3) the elasticity of corruption to the tax rate: $\epsilon^{u}_{\tau,i} = \frac{u_i^{'}(\tau_i)}{u_i(\tau_i) - u_i(\bar{\tau}_i)} (\tau_i - \bar{\tau}_i)$.

The optimal effective tax rates in each industry are characterized by the optimal conditions of the maximization problem. To simplify the analysis, I consider the case when θ is equal to zero. That is, the tax administrator can only obtain a negligible fraction of rent as rewards. He/she would not intentionally raise the tax rate to force more firms to bribe. This is true if the bribing firms resort to the more powerful politicians for help rather than finding help from the tax administrator. Therefore, most, or even all, bribes may not go into the tax administrator's pocket.

The optimal tax structure when $\theta = 0$ can be characterized by the proposition as below. Of course, the proposition still holds when θ is sufficiently small.

Proposition 1. *(The Tax Rate Distribution across Industries) If $\theta = 0$, then the optimal effective tax rate in each industry is characterized by $B'(T) = G_i(\tau_i)C_i^{'}(T_i)$ for any $i =$*

$1, 2, ..., N$, *where* $G_i(\tau_i) = 1 - \frac{1-1/\epsilon^c_{\tau,i}}{1-\epsilon^T_{\tau,i}/\left|\epsilon^u_{\tau,i}\right|}$ *and* $G_i(\tau_i)C_i'(T_i)$ *is the adjusted marginal tax collection cost.*

(1) Without corruption, that is, $\epsilon^u_{\tau,i} = 0$*, then* $G_i(\tau_i) = 1$*, and the optimal effective tax rates should equate to the marginal tax collection cost* $C_i'(T_i)$ *across industries.*

(2) With corruption, that is, $\epsilon^u_{\tau,i} < 0$*, then* $G_i(\tau_i) > 1$*, and the the optimal effective tax rates should equate to the adjusted marginal tax collection cost* $G_i(\tau_i)C_i'(T_i)$ *across industries.*

The following proposition demonstrates, in a simplified setting, that the change of $\left|\epsilon^u_{\tau,i}\right|$ in one industry can affect the whole distribution of the effective tax rate across industries.

Proposition 2. *(Change of Tax Rate Distribution across Industries) Suppose* T *does not change (or* $B(T)$ *curve is vertical and* $B'(T) = 0$ *) and* $\tilde{T}'(\tau_i) > 0$ *. If* $\left|\epsilon^u_{\tau,i}\right|$ *in industry* i *increases as political and business environment becomes more corrupt, then the effective tax rate should fall in this industry and rise in other industries.*

The proof of this proposition is simple. We know that $G_i(\tau_i)C_i'(T_i)$ is an increasing function in τ_i and $\frac{\partial G_i(\tau_i)}{\partial\left|\epsilon^u_{\tau,i}\right|} > 0$. Suppose $\epsilon^c_{\tau,i}$ and $\epsilon^T_{\tau,i}$ are constant and determined by technology and market structure. $\left|\epsilon^u_{\tau,i}\right|$ may change with the business environment. Given the optimality condition $G_i(\tau_i)C_i'(\tau_i Y(\tau_i)) = 0$, the increase of $\left|\epsilon^u_{\tau,i}\right|$ shifts the curve $G_i(\tau_i)C_i'(\tau_i Y(\tau_i))$ upward and thus reduces the optimal τ_i. The total tax revenue $\tilde{T}(\tau_i)$ consequently falls for industry i. Since T is fixed, the tax revenue for other industry $\tilde{T}(\tau_{-i})$ should increase, which requires τ_{-i} to rise for other industries. In the case of $B'(T) < 0$, the proof is a bit more complicated but the conclusion still holds under some broad conditions.

Proposition 2 suggests that the effective VAT rate should rise for capital-intensive industries over the tenure of prefectural secretary while it drops for labour-intensive industries, if we believe the business environment becomes more corrupt as the tenure of a prefectural secretary rises, then the corruption of some firms would create the general equilibrium effect on other un-corrupt firms, and, if the capital-intensive firms tend to be more corrupt, that is, their $\left|\epsilon^u_{\tau,i}\right|$ rises over the tenure, then we would observe the fall of VAT rate in capital-intensive industries and its rise in labour-intensive industries.

2.2 Bribing Model and the Structural Form of ϵ^u_τ

The optimal tax condition suggests that the cross-industry distribution of the effective tax rate relies heavily on $\left|\epsilon^u_{\tau,i}\right|$, the elasticity of corruption to the tax rate in each industry. A question still remains unanswered. Given the business environment, which tends to offer equal corruption opportunities to all firms, why are capital-intensive firms more susceptible to corruption than their labour-intensive counterparts? Why is the level of corruption of capital-intensive firms more sensitive to the effective tax rate?

The bribing model in this part attempts to answer these questions and to provide a structural form for ϵ_{τ}^{u}. The model has two features: (1) fixed corruption cost; and (2) Nash bargaining over the corruption rent. It can generate the complementarity between tax benefit and other benefits from corruption, which is key to answering the questions.

For narrative convenience, the firm size Y_i is normalized to one in this model. Suppose the firms j in each industry i now are heterogeneous in that they need to pay a firm-specific fixed cost s_{ij} before being able to bribe. s_{ij} is like a club membership fee in order to enjoy the corruption service. Within each industry, s_{ij} satisfies a distribution represented by CDF function $F(s)$ over the support $[0, \infty)$. The distribution is the same for all industries.

There are two kinds of rents created from corruption: preferential tax treatment $\tau_i - \bar{\tau}_i$ and other benefits B_i. The latter may include easy access to finance, cheap land, and so on. The total rent is $\tau_i - \bar{\tau}_i + B_i$.

The rent is split between bribing firms and corrupt officials under the Nash bargain, with the firm having bargaining power $\beta \in (0, 1)$. If a firm has decided to bribe, then it needs to pay a bribe $Q_i = \beta (\tau_i - \bar{\tau}_i + B_i)$ to the official who has promised to help. So the firm's benefit from corruption is $R_i = (1 - \beta)(\tau_i - \bar{\tau}_i + B_i)$.

The decision of a firm whether or not to participate in corruption is made in the following way. The firm first observes its s_{ij} and R_i. It would like to bribe if $s_{ij} < R_i$. In this case, it need to pay both the fixed cost s_{ij} and the bribe Q_i. Otherwise, it would participate in corruption.

Under these settings, the elasticity of corruption to the tax rate ϵ_{τ}^{u} can be pinned down as the following structural form.

Proposition 3. *(Structural Form of the Elasticity of Corruption to Tax Rate)* $\epsilon_{\tau}^{u} = \lambda_i (\tau_i - \bar{\tau}_i)$, *where* $\lambda_i = \beta \frac{f(R_i)}{1 - F(R_i)}$, *and* $R_i = (1 - \beta)(\tau_i - \bar{\tau}_i + B_i)$. ϵ_{τ}^{u} *satisfies* $\frac{\partial \left| \epsilon_{\tau,i}^{u} \right|}{\partial B_i} > 0$ *if the hazard rate* $\frac{f(R_i)}{1 - F(R_i)}$ *is rising in* R_i *(or* s_{ij} *has a thin-tailed distribution).*

Intuitively, this proposition suggests the complementarity between the tax benefits and other benefits from corruption. The bigger the other benefits B_i, the more sensitive corruption is to the effective tax rate.

Empirically, the paper has shown that capital-intensive firms generally obtain more other benefits, such as access to finance. Bigger B_i implies bigger $\left| \epsilon_{\tau,i}^{u} \right|$. Combined with Proposition 2, it should lead to lower τ_i in capital-intensive industries and higher τ_{-i} in labour-intensive industries. As the tenure of the party secretaries rises, more corruption opportunities are offered. Capital-intensive firms are more liable to participate in corruption and more sensitive to τ_i. Therefore, τ_i falls for capital-intensive firms and τ_{-i} rises for labour-intensive ones.

3 Institutional Background

Local Governance in China and the Communist Party Secretary

China is ruled under a five-level hierarchical structure of governments – central, province, prefecture, county, and village/town. At each level of government, the *de facto* political leader is the secretary of the Communist Party of China (CPC). At the prefecture level, the prefectural secretary of the CPC outranks the mayor, legislators, and judges in most of the major local government decisions. The political power of party secretaries originates from their right to appoint personnel.[6]

The governance of China is centralized through the personnel appointment and control of Communist Party secretaries at the different levels. Usually, the prefectural secretary is appointed by the Provincial Committee of the CPC. In most cases, the timing of this appointment coincides with the Prefectural Congress of Party Representatives, which takes place every five years and therefore determines a standard five-year term for prefectural secretaries.[7]

Party Secretary and VAT Administration

Although VAT is administered by the State Administration of Taxation and the appointment of its personnel and funding allocations are under the vertical administration of the central government, a local party secretary can still exert enormous influence, either directly or indirectly, on VAT inspectors.

Within the system of the Communist Party of China, a prefectural party secretary is the direct supervisor of the party secretary of the State Administration of Taxation of the same prefecture. The former can considerably affect the promotion prospects and political career of the latter. In almost all cases, the chief of the State Administration of Taxation is also the party secretary of the State Administration of Taxation, and is therefore subject to the supervision of the prefectural party secretary.

In addition to the direct supervisory role of prefectural party secretaries over the chief of the SAT, there are several other indirect ways through which local governments and secretaries can intervene in VAT enforcement. First, the local government can help the SAT in issues such as obtaining land for office buildings, schooling for children, local hospitals for health care, and so on. Second, it is openly known that the SAT receives a subsidy from the local government in order to improve its working conditions and staff

[6] One of the political principles in China is that the cadre should be under the rule of the Party (CPC) (or "Dang Guan Gan Bu" in Chinese. A cadre is generally a party member in key position at different level of governments). The power of secretaries of the CPC, originated from personnel appointment, is known as the "the mother of power". (*People's Forum* (*Renmin Luntan*), issue 2, 2007). For more introduction to secretaries of the CPC, refer to "http://baike.baidu.com/view/2372141.htm?func=retitle".

[7] Xu (2011) offers an insightful and comprehensive description of the characteristics of China's political regime.

welfare, or for any other reason. Third, the capability of the SAT in tax law enforcement is limited by local departments of Public Security, which are under the tight control of local governments.[8] Fourth, tax administrators have their own dirty laundry. Selective anti-corruption is a popular threat of local governments against a tax administrator once he/she refuses to cooperate.

There are many potential ways through which local politicians can influence the *de facto* tax rate. Here I introduce two possible ones. First, local governments can demand that the SAT turns a blind eye to firms using fake invoices, which are rampant across China.[9] Second, local governments can directly order the local office of the SAT to enforce a lower VAT rate on a firm, if the local officials are powerful enough.[10]

To sum up, although the *de jure* VAT rates are legislated at the national level, prefectural party secretaries and other local officials can influence the *de facto* tax rate of a firm.

4 Data and Main Variables of Interest

4.1 Data

The paper employs four datasets: (1) the Annual Survey of Industrial Production conducted by the National Bureau of Statistics of China (2000-2007); (2) newly-digitized data on political turnovers of prefectural secretary of the Chinese Communist Party (2000-2010); (3) the Investment Climate Survey conducted by the World Bank in China (2005); (4) the County Public Finance Statistics Yearbook of China (2000-2007).

The Annual Survey of Industrial Production includes all state-owned firms and non-state owned firms with annual sales of more than 5 million RMB yuan (approximately 800,000 US dollars). I use the data from 2000 to 2007. Number of firms increases from

[8]For example, in China, the SAT is entitled to crack down on fake invoices only with the help of the police. Therefore, the SAT can do nothing with a fake invoice if the local police are not willing to collaborate.

[9]Anecdotal evidence show that this type of case is not rare in many regions. A case is currently being brought to the court in Jiang Xi province in which a company is being charged with issuing and selling fake VAT special invoices to other firms. The company under charge is said to be subsidised by the local government. For another example, the SAT annouced the eight biggest cases of tax law violation in 2013. All of these cases were related to fake VAT special invoices. In every case, the tax evaded is above one hundred million RMB yuan (about 15 million US dollars). The highest effective tax rate the firm paid in one of these cases was only 0.11% (Xinhua Net, October 21, 2013. http://news.xinhuanet.com/2013-10/21/c_117804571.htm)

[10]For example, anecdotal evidence reveals that in 2012, Samsung Electronics decided to invest a project worth seven billion US dollars in Xi'an, capital city of Shannxi Province. To attract the gigantic FDI, Mr. Zhao, the provincial secretary of the CPC at the time, agreed that a VAT rate of 11% could be applied to this project. Of course, disputes over the unlawful tax rate followed between the Shannxi government and the SAT. The central government had to step in and finally Shannxi government won the case. It turned out to be a gentleman's agreement through private bargaining between the local government, the SAT and the cetral government.

about 66,000 to 168,000 during the sample period after dropping bad observations. Information on each firm includes a 4-digit industry code, ownership, county-level region code, value-added, sales revenue, and tax payments including VAT, corporate income tax, business tax and other minor local taxes.

Data on the political turnover of prefectural secretaries of the Chinese Communist Party is collated from government websites. The dataset is a balanced panel containing all prefectures (there are around 300 prefectures in China) and all years from 2000 to 2010. The data reveal whether there is a change in prefectural secretary of the CPC for each prefecture in each year and, if there is a change, whether the new secretary is locally promoted or appointed from elsewhere.[11]

The Investment Climate Survey Data in China conducted by the World Bank (2005) includes 12,000 manufacturing and service sector firms in 120 prefectures of China. The questionnaire covers a huge range of information regarding the investment climate. In this paper, we are interested in only two types of firm records: (1) tax payments and balance sheets; and (2) the relationship between firms and governments.

The County Public Finance Statistics Yearbook of China (2000 - 2007) includes government revenue and government expenditure at the county and prefecture levels. In this paper, we are interested in the following variables: (1) tax revenue; (2) total revenue (including off-budget revenue); (3) total expenditure (including off-budget expenditure); (4) agricultural taxation revenue; (5) subsidy for the agricultural taxation reform; (6) total population; and (7) total GDP.

All the summary statistics of main variables are reported in Table 1. It should be noted that the firm-level data in the Annual Survey of Industrial Production are collapsed at the county-(2-digit)industry-year level for regressions. Similarly, the data in the Investment Climate Survey prefecture-(2-digit) level are collapsed at the prefecture-(2-digit)industry level because it does not provide the county code.

Table 1: Summary Statistics (INSERT)

4.2 Measurement of Main Variables

4.2.1 Effective VAT Rate

I construct the effective VAT rate of a firm f in year t as:

$$Effective\ VAT\ rate_{f,t} = \frac{Payble\ VAT_{f,t}}{Value\ added_{f,t}} \tag{4.1}$$

[11] Of course, one may propose using even smaller county-level jurisdiction to do the same work. But the problem with the county level is there are about three thousand county-level jurisdictions in China. On average, there will be only three firms in each country-(2-digit)industry-year cell, leaving the study contaminated by a potentially large sampling error.

which can be re-written as $\frac{\tilde{\tau}^s_{f,t}\cdot\tilde{S}_{f,t}-\tilde{\tau}^m_{f,t}\cdot\tilde{M}_{f,t}}{S_{f,t}-M_{f,t}}$, where $\tilde{S}_{f,t}$ and $\tilde{M}_{f,t}$ are the sales and intermediate inputs used to calculate the payable VAT. For notational convenience, I ignore subscripts f and t for all relevant variables in this section. These are the numbers recorded by firms on their VAT special invoices. $\tilde{\tau}^s$ and $\tilde{\tau}^m$ are the tax rates actually applied for sales and intermediate inputs, respectively. They could differ from the statutory tax rates τ^s and τ^m. In the data, only $\tilde{\tau}^s \cdot \tilde{S}$ and $\tilde{\tau}^m \cdot \tilde{M}$ are observable.

To understand sources of variation in the effective VAT rate, I decompose it into two components as:

$$Effective\,VAT\,rate = \frac{\tilde{\tau}^s\cdot\tilde{S}-\tilde{\tau}^m\cdot\tilde{M}}{\tau^s\cdot S-\tau^m\cdot M}\cdot\frac{\tau^s\cdot S-\tau^m\cdot M}{S-M} \tag{4.2}$$

The second component on the right-hand side of Expression (4.2), $\frac{\tau^s\cdot S-\tau^m\cdot M}{S-M}$, can be re-written as $\tau^s + (\tau^s-\tau^m)(S/M-1)^{-1}$, which implies that the variation in the statutory rates τ^s and τ^m, as well as input-output structure S/M, could potentially be sources of variation in the effective VAT rate. If the firm exports goods worth E, then τ^s in Expression (4.2) should be replaced by $\tau^s - \tau^e \cdot (E/S)$, where τ^e is the post-rebate statutory VAT rate for exports. In this case, the ratio of export to sales is an additional source of variation.[12] Although the second component contains various possible sources of variation described above, Chen (2015a) shows that they are insufficient to explain variation in the effective VAT rate. Therefore, this paper focuses on the first component and investigates the variation of the effective VAT rate as a result of lax tax enforcement or tax evasion.

Several caveats should be noted with the measurement of the effective VAT rate. First, because only $\tau^s \cdot \tilde{S}$ and $\tau^m \cdot \tilde{M}$ are observable, we cannot distinguish the exact ways of tax evasion by manipulating the tax rate (τ^m or τ^s) and by the tax base ($\tilde{M}$ or $\tilde{S}$). Second, one may suspect the variation of the effective VAT rate may come from the manipulation of reported value-added, that is, the denominator in Expression (4.1), rather than from tax evasion through the numerator.

4.2.2 Tenure of Prefectural Party Secretary

There are two variables regarding the tenure of prefectural secretaries of the CPC: (1) tenure of secretary; and (2) duration of secretary in office. The tenure of the secretary measures the years that the secretary has been in office since he/she was appointed. The duration in office indicates the total years for which the secretary is in office between his appointment and the appointment of his successor.

Figure 4.1 describes the turnover of prefectural secretaries of the CPC from 2001 to 2010 in China. The figure shows that the turnover is quite random over time. Around 80

[12]The staturory rate varies because certain sectors or activities are taxed at different rates, and firms differ in their input-output structure or in the extent to which they engage in activities subject to special tax treatment.

out of 300 prefectures undergo political turnover each year.

Figure 4.1: Turnover of Prefectural Secretary of CPC

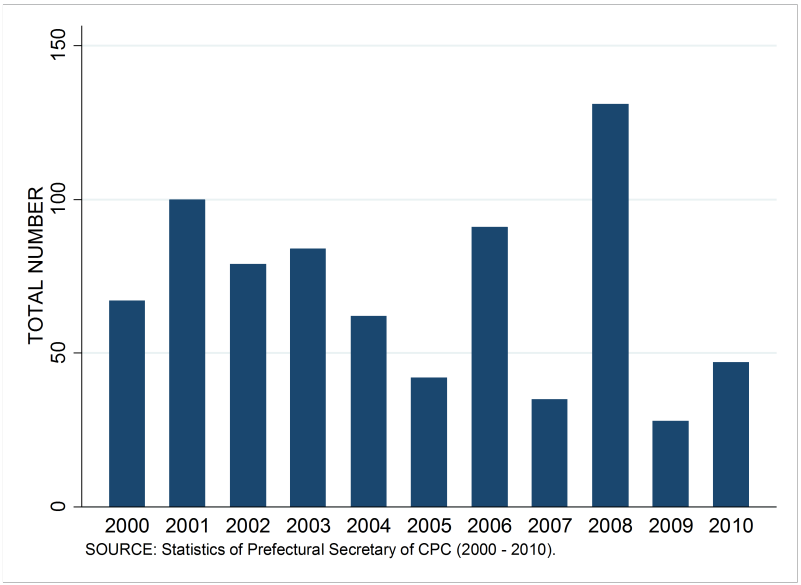
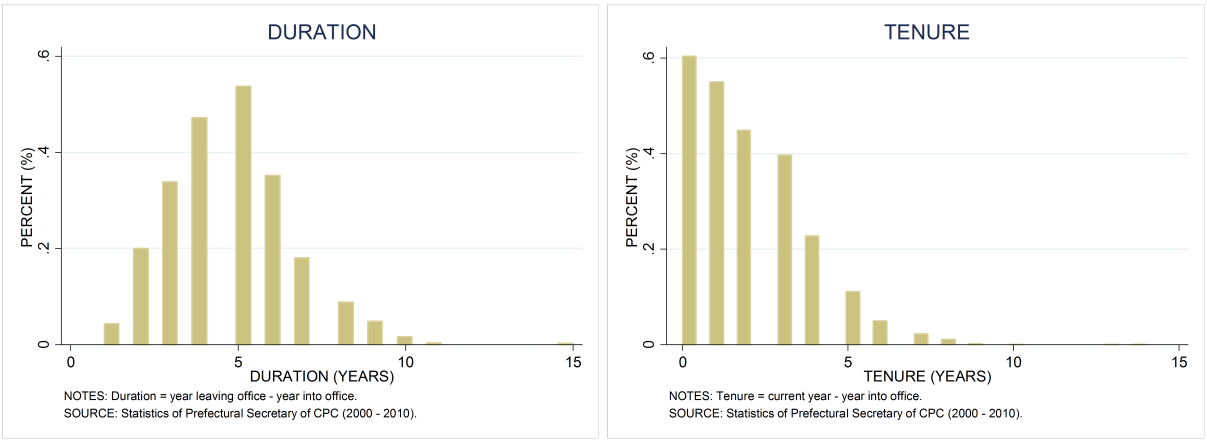


Figure 4.2 depicts the distribution of tenure and duration of prefectural secretaries in the whole of China. From the distribution of duration (left graph) we can see that secretaries are most likely to leave office after five years, which is one standard term in office for the secretary.

Figure 4.2: Distribution of Tenure and Duration of Prefectural Secretary of CPC



5 Party Secretary's Tenure and the Distribution of Effective VAT Rate across Industries

5.1 Changes in VAT Rate over Secretary Tenure

Motivated by Figure 4.2, this paper aims to uncover the political factors that shape the distribution of the effective VAT rate. In particular, the paper shows that the cross-industry distribution of the effective VAT rate and capital intensity changes over the tenure of prefectural secretary.

Figure 5.1: Tenure and Variation of Effective VAT Rate with Capital Intensity

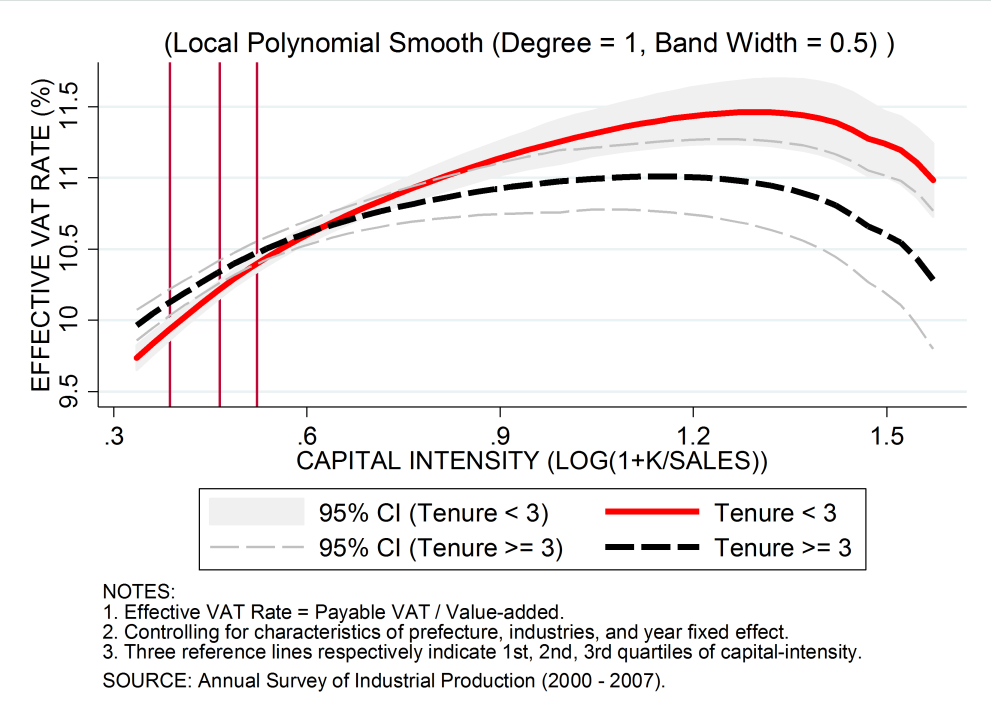


Figure 5.1 displays how the variation of the effective VAT rate with capital intensity changes over the tenure. The capital intensity is defined as $LOG(1 + total\ assets/sales)$. The figure clearly shows that the effective VAT rates of capital-intensive firms go down as the tenure rises, while the effective VAT rates of labour-intensive firms go up.

Figure 5.1 is drawn using the following steps. First, firms are divided into two groups according to the tenure of the prefectural secretary in the corresponding region and year. Because a standard term is five years, observations are classified as one group if tenure < 3, another for tenure ≥ 3. The residual effective VAT rate can then be obtained by a regression on a list of relevant factors.[13] Finally, the variation of the (residual) effective

[13]These factors include: (1) firm characteristics: exporst/sales, input-output structure measured by intermediate-input/output, ownership type, profitability measured by profit/sales, firm size measured by log(value-added), log(sales), log(total asset); (2) prefecture-(1-digit)industry fixed effect.

VAT rate with capital intensity for each group is drawn using a local polynomial smooth method (degree = 1, bin = 0.5). Three red vertical reference lines indicate the quartiles of capital intensity across industries. The figure shows that fewer than a quarter of industries experience a reduction in their VAT rate, while the remaining three-quarters experience their effective VAT rate rising.

5.1.1 Empirical Strategy

Motivated by Figure 5.1, the main regression is specified as below to capture the heterogeneous tenure effect on the relationship between the effective VAT rate and capital intensity.

$$\begin{aligned}\tau_{c,i,t} = \alpha \ + \ \eta_{p,s} \ + \ \mu_{r,t} \ + \ \gamma \cdot CapInt_i \cdot Year_t \ + \\ \lambda \cdot Tenure_{p,t} \ + \ \varphi \cdot Tenure_{p,t} \cdot CapInt_i \ + \\ \rho \cdot CapInt_i \ + \ \psi \cdot CapInt_i^2 \ + \ \theta \cdot X_{p,i} \ + \ \beta \cdot Tenure_p \cdot X_{p,i} \ + \\ \delta \cdot Z_{c,t} \ + \ \sum_{d=1}^{3} \omega_d \cdot Duration\_Dummy_{p,t}^{d} \ + \ \epsilon_{c,i,t}\end{aligned} \tag{5.1}$$

where the outcome variable $\tau_{c,i,t}$ is the mean of the "effective VAT rate" within a country-(2-digit)industry-year cell. Subscripts c, p, r, i, s, t represent county, prefecture, province, 2-digit industry, 1-digit industry and year, respectively.

Key variables on the right-hand side are $Tenure_{p,t}$ and $CapInt_i \cdot Tenure_{p,t}$, where $Tenure_{p,t}$ is the tenure of the secretary of the CPC in prefecture p in year t, which takes natural numbers {0, 1, 2, 3, ...} in main regressions. $CapInt_i$ is defined as in Figure 5.1, except that I subtract its minimum value across 2-digit industries, so that $CapInt_i = 0$ for the least capital intensive industry.[14]

There are three region-industry-time specific controls. $\eta_{p,s}$ is a prefecture-(1-digit)industry fixed effect, which captures the persistent region-industry factors determining $\tau_{c,i,t}$. $\mu_{r,t}$ is the province-year fixed effect, which absorbs the province-year specific shocks, such as political turnover at the province or the national level. $CapInt_i \cdot Year_t$ is to absorb the industrial trend which depends on capital intensity.

To control for the tenure effect on taxation through characteristics other than capital intensity, I add $Tenure_{p,t} \cdot X_{p,i}$ in the regression, where $X_{p,i}$ is a vector of other firm characteristics of 2-digit industry i in prefecture p over all years, including ownership, profitability, firm age, logarithm of sales volume, logarithm of value-added, ratio of intermediate input to output, and ratio of exports to sales. To avoid the endogeneity problem, following Fisman and Svensson (2007), all these firm characteristics, except ownership,

[14]This paper finds that capital intensity is an important industry characteristics according to which the tenure effect shows heterogeneity. The problem is that, conditional on sales volume, capital-intensive firms are also the firms with the greatest capital. One way to disentangle the concept of capital intensity and firm size is to use investment intensity rather than capital intensity. The paper performs a robustness check with this replacement, and the results are consistent.

take the median within 2-digit-industries of prefecture p (while ownership takes the mean due to its 0-1 nature). [15] In addition to $CapInt_i$, its square term $CapInt_i^2$ is also included to capture the non-linear relationship between $\tau_{c,i,t}$ and $CapInt_i$ shown in Figure 5.1.

$Z_{c,t}$ is a vector of county characteristics, including logarithm of GDP per capita, logarithm of total population, and fiscal budget pressure faced by the county-level government.[16] These may affect the incentives for tax enforcement and the taxation capacity of local governments. Inclusion of local GDP is also due to the concern over the manipulation of GDP data by local officials. This could affect the effective VAT rate if a firm's tax payment is not adjusted, together with the manipulation of value-added data.[17]

I also divide prefectual secretaries into three groups, and add three dummy variables $Duration\_Dummy_{p,t}^d$ $(d = 1, 2, 3)$, according to whether their *ex post* duration is less than, equal to, or more than five years (the standard duration of one term), respectively. I do this because less capable secretaries may find it harder to get promoted and tend to stay in office longer than one term. The estimates of tenure effect will be biased if these personal characteristics affect their incentives and ability to intervene in tax enforcement.

To show the dynamics of tenure effect, I also present results where $Tenure_{p,t}$ is replaced by a vector of dummy variables, each for a different possible number of years of tenure. Because fewer than 8% of the observations have a tenure longer than five years, the standard error could be large for these observations. In the empirical implementation, observations with tenure longer than five years will be regarded the same as those with tenure of five years.

5.1.2 Results

The first two rows in Table 2 show the results of the tenure effect on firms' effective VAT rates based on Equation (5.1).

Panel A does not include $Tenure_{p,t} \cdot X_{p,i}$ in the regression. Column (1) shows no significant tenure effect on the effective VAT rate across industries if we do not consider heterogeneity based on capital intensity. Column (2), however, reveals significant opposing tenure effects on labour-intensive and capital-intensive industries. The coefficients imply that as the tenure increases, the effective VAT rate increases for the most labour-intensive

[15] For example, measurement errors of explanatory variables can generally result in biased estimation. In regression (5.1), the mean of capital-intensity of an industry is sensitive to the extreme measurement error of outliers, but its median is not.

[16] Fiscal budget pressure for the county government is measured by: [total budget expenditure - basic construction expenditure - total budget revenue - (extra-budget revenue - extra-budget expenditure)] / total population.

[17] One caveat should be noted. Due to the lack of data on county characteristics for some counties or in some years, the regression sample size would reduce by half once county characteristics are controlled for. However, the number of prefectures covered in the sample would not dramatically drop as most prefectures would still have some counties in the sample. This is not a fatal problem as my interest is the variation within-prefecture and over time.

industries and decreases for the more capital intensive ones. Column (3) reports the key results that control for industrial trends with "$CapInt * Year$", and shows that the tenure effects are still significant. However, their magnitude are absorbed slightly by the industrial trends.

Panel B includes the cross-terms $Tenure_p \cdot X_{p,i}$. The coefficients of "$Tenure$" and "$Tenure * CapInt$" in all four columns confirm the baseline results in Panel A. Additionally, the results of "*Tenure * other industry characteristics*" in Columns (4) – (6) show that other industry characteristics are not the channels through which tenure affects the effective VAT rate. Columns (1) - (4) all control for the county characteristics, including LOG(GDP/Pop), LOG(Pop), and fiscal pressure. In Columns (5) and (6), I drop these three controls. This allows me to use a larger sample including counties that do not provide data on these three controls. As a remedy, Column (6) replaces the prefecture-(1-digit)industry fixed effect with a county-(1-digit)industry fixed effect. Compared to columns (5) and (6), the significance of "$Tenure$" goes up while that of "$Tenure * CapInt$" goes down.

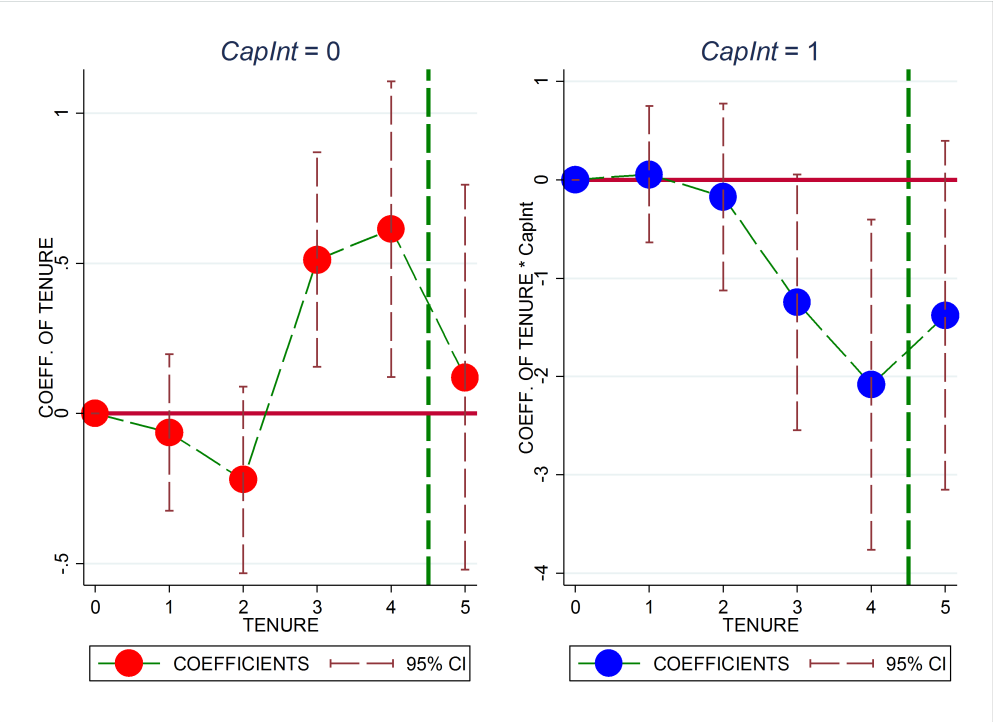
The results in Column (6) bear some quantitative implications, suggesting that when the prefectural secretary stays in office for additional year, the effective VAT rate rises by 8% (or 0.8 percentage point from 9.5% to 10.3%) for labour-intensive industry with $CapInt = 0$, and drops by 3% (or 0.3 percentage point from 11% to 10.7%) for capital-intensive industry with $CapInt = 1$. Given the size of the economy involved, the magnitude are not negligible.

Table 2: Tenure Effect on Effective VAT Rate (INSERT)

Figure 5.2 shows the dynamic heterogeneous tenure effect on the variation of the effective VAT rate with capital intensity. The tenure effect is still estimated based on Equation (5.1) and the regression follows the version of column (5) in Table 2, but with $Tenure_{p,t}$ replaced by a vector of dummy variables $\left\{Tenure^s_{p,t}\right\}^D_{s=0}$. The left panel shows the tenure effect on for the most labour-intensive industry, namely the one with $CapInt_i = 0$. The right panel shows the tenure effect on a more capital intensive industry ($CapInt_i = 1$), which corresponds to the $98th$ percentile of the distribution of $CapInt_i$. Several features of Figure 5.2 should be noted. First, the tenure effect is not significantly different from zero at the 95% significance level in the first two years for either labour-intensive industries or capital-intensive industries. Second, in year three, the tenure effect is significant but moves in the opposite direction for labour-intensive and capital-intensive industries – positive for the former and negative for the latter. Third, the effects remain at around the same level until at least year five, when the prefectural secretary finishes his/her standard first term and is quite likely to leave office. [18]

[18] The cyclical pattern of tenure effect on the effective VAT rate with peak in year four might be possibly

Figure 5.2: Tenure Effect on Variation of Effective VAT Rate with Capital Intensity



5.2 Tenure-specific Response to a Natural Experiment

In Section 5.1, we have focused on the effect of the prefectural secretary of the CPC's tenure on the LEVEL of the effective VAT rate. In this section, I study the effect of the tenure on the CHANGE in the effective VAT rate across industries in a situation where local governments have to strengthen tax enforcement for all industries.

Using a Diff-in-Diff approach, Chen (2015b) finds that, after the abolition of taxes on agricultural activities, effective VAT rates increased disproportionately in counties that had previously been most reliant on these taxes. This finding suggests that abolition of the agricultural tax is a good natural experiment in which the average level of tax enforcement is strengthened in response to public revenue losses.

Combining the heterogeneous tenure effect on the effective VAT rate discussed in Section 5.1 and the change of tax enforcement due to the abolition of agricultural tax, this section attempts to answer the following questions. After the abolition of agricultural tax, is the tenure effect on the CHANGE of the variation of the effective VAT rate with capital intensity consistent with that on the LEVEL of the effective VAT rate? Does the

explained by the capability and incentive of corruption over the local political cycles. Given a five-year average term of prefectural secretaries, the window for corruption is narrow for local officials. In the first two years, the new political network is yet to be in shape. Corruption is hard to carry out. In year five, when prefectural secretaries are most likely to get promoted, corruption may hazardously wreck their political career. Therefore, corruption are most likely to take place in year three and year four in the middle of a term.

CHANGE of the effective VAT rate also favour capital-intensive firms more in the region where the prefectural secretary stays in office for longer?

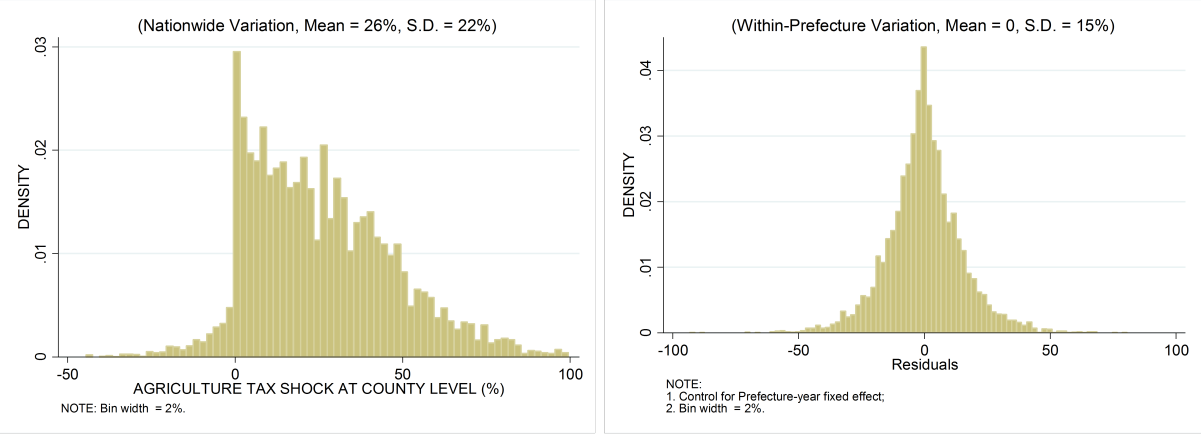
To see how the abolition of agricultural taxation reduced local tax revenue and how this affected local governments' incentives to strengthen tax enforcement, this paper follows Chen (2015b) and measures the agricultural taxation revenue shock for each county c as:

$$Agr_c = \frac{(Agr\ tax\ revenue_{c,2000\sim2004} + Subsidy_{c,2000\sim2004})}{total\ tax\ revenue_{c,2000\sim2004}} - \frac{Subsidy_{c,2005\sim2007}}{total\ tax\ revenue_{c,2005\sim2007}} \tag{5.2}$$

On the right-hand side, $X_{c,2000\sim2004}$ ($X_{c,2005\sim2007}$) is the average of variable X in county c between 2000 and 2004 (2005 and 2007). *Agr tax revenue* is the agricultural tax revenue. *Subsidy* is the central government transfer associated with the agricultural taxation reform and received at the county level. To avoid bargaining between central government and local government, the subsidy is based on a pre-determined formula. *total tax revenue* is the total budgetary tax revenue, including VAT, corporate income tax, business tax and others, while off-budget revenue is not included.[19]

Figure 5.3 plots the distribution of agricultural reform shocks across approximately 3,000 county-level jurisdictions in China. The left graph is the original value of the shock calculated based on Expression (5.2); the average shock is 25%. It should be noted that the shocks to a small number of counties are negative, showing that these counties could even have benefitted from the reform as they received more in subsidy than they lost in taxes following the reform. The right graph, by normalizing the prefecture average shock to zero, displays the shock across counties within a prefecture. The shock is quite symmetric across counties, with standard deviation at about 15%.

Figure 5.3: Distribution of Agriculture Tax Shock across Counties



The nationwide abolition of agricultural tax generates variation of tax revenue loss across two dimensions. Over time, most of the regions suffered a negative revenue loss,

[19] In China, the local government can also raise revenue through other non-tax sources such as fees and local funds, or selling land and public assets such as state-owned enterprises.

while across regions, the intensity of the shock varied. This variation over time and across regions allows for the standard Diff-in-Diff method to identify the impact on tax enforcement. Chen (2015b) uses the Diff-in-Diff method to demonstrate that the treatment effect of the abolition of agricultural tax on the effective VAT rate is positive, and the parallel trend assumption is satisfied.

Based on the Diff-in-Diff method, the heterogeneous tenure effect on the variation of tax enforcement with capital intensity can be captured by the following regression.

$$\begin{aligned}\tau_{c,i,t} = \alpha \ + \ \beta_{c,i} \ + \ \eta_{p,s,t} \ + \ \gamma \cdot Post_t \ + \\ (\lambda \cdot Tenure_{p,t} \ + \ \varphi \cdot Tenure_{p,t} \cdot CapInt_{r,i,t}) \cdot Post_t \cdot Agr_c \ + \\ \rho \cdot X_{c,i,t} \ + \ \epsilon_{c,i,t}\end{aligned} \tag{5.3}$$

$Post_t$ is the dummy variable for years *post* the agricultural tax reform. Agr_c, measured by expression (5.2), is the agricultural reform shock for county c, and the vector $X_{c,i,t}$ contains all the levels and interactions of terms between $Post_t$, $CapInt_{r,i,t}$, and Agr_c, as well as additional controls I used in Equation (5.1).

The coefficients λ and φ tell us how the distribution of the burden from local governments having to respond to a fiscal shortfall varies with the CPC secretary's tenure. A positive λ and negative φ imply that, the longer the secretary has been in office, the more the burden shifts to labour-intensive industries.

I include a prefecture-(1-digit)industry-year fixed effect $\eta_{p,s,t}$ to absorb the prefecture-(1-digit)industry-year specific shock potentially correlated with the county agricultural tax shock or turnover in party secretary shock at the prefecture level. I also include a county-(2-digit)industry fixed effect $\beta_{c,i}$ to control for the time-invariant county-industry characteristics do not need to be included in the regression.

It should be noted that firm characteristics at the (2-digit)industry level should not be included in the regression because $\beta_{c,i}$ has been introduced. Capital intensity $CapInt_{r,i,t}$ is the median at the province-(2-digit)industry-year level so that it is not affected by the prefecture-(2-digit)industry-year shock. Other firm characteristics take median at the county-(2-digit)industry-year level (ownership takes the mean due to its 0-1 nature), to control for their impact on $\tau_{c,i,t}$. The same as Equation (5.1), $\sum_{d=1}^{3} \omega_d \cdot Duration\_Dummy_{p,t}^{d}$ is introduced to avoid the selection problem.

5.2.1 Results

Based on the regression of Equation (5.3) with the firm-level data collapsed at the county-year level, Table 3 reports the heterogeneous tenure effect on the variation of VAT enforcement triggered by the abolition of agricultural tax with capital intensity.

Panel A reports the Diff-in-Diff baseline regression results. The treatment effect of agricultural tax reform on VAT enforcement is captured by the coefficient on "$Post * Agr$".

To cope with the within-group serial correlation problem, as identified by Bertrand et al. (2004), robustness standard errors are clustered at the prefecture level in column (1). As a robustness check, clustering are at the province level in column (2).

Panel B reports the regressions based on Equation (5.3). The key results are the coefficients on "*Post * Agr * Tenure*" and "*Post * Agr * CapInt * Tenure*". They show that after the agricultural tax reform, the counties suffering greater agricultural tax revenue shock strengthened their VAT enforcement on all industries, but with no difference between industries (coefficient of "*Post * Agr * CapInt*", $\mu = 0$). However, in counties where prefectural secretaries were in office longer, the enforcement is higher on labour-intensive industries (coefficient of "*Post * Agr * Tenure*", $\lambda > 0$) and lower on capital-intensive industries (coefficient of "*Post * Agr * Tenure * CapInt*", $\varphi < 0$). Robustness standard errors are clustered at the province-(1)industry level in Panel B.

Table 3: Tenure Effect on the Change of Tax Enforcement (INSERT)

Figure 5.4: Tenure Dynamic Effect on Change of VAT Rate

Figure 5.4 displays the dynamic tenure effect based on regression (5.3), with $Tenure_{p,t}$ replaced by a vector of dummy variables for each tenure.[20] The results are consistent

[20]Similar to the approach in Section 4, $\lambda \cdot Tenure_{p,t} + \varphi \cdot Tenure_{p,t} \cdot CapInt_{r,i,t}$ in Equation (5.3) is replaced with $\sum_{s=0}^{D} \left(\lambda_s \cdot Tenure_{p,t}^{s} + \varphi_s \cdot Tenure_{p,t}^{s} \cdot CapInt_{r,i,t}\right)$, and other terms involving $Tenure_{p,t}$ do not change to avoid unnecessary complications. The dynamics of the tenure effect for labour-intensive and capital-intensive industries can be caputured by $\{\lambda_s\}_{s=0}^{D}$ and $\{\varphi_s\}_{s=0}^{D}$, respectively.

with Table 3 where $Tenure_{p,t}$ takes the natural measurement of tenure. The effective VAT rate rises for the labour-intensive industries (left panel) and falls for the capital-intensive industries (right panel) as the tenure of the prefectural secretary increases. Once again, the results confirm the rising favouritism towards capital-intensive industries as the prefectural secretary stays in office longer.[21]

6 Other Tenure-sensitive Outcomes

In this section, I study the effect of tenure on the misreporting of profit. As Cai and Liu (2009) suggested, profit misreporting is potentially related to the evasion of corporate income tax.

Additionally, it is well known that credit allocation in China is subject to political influence (Li et al., 2008; Allen et al. 2005; Ponet et al., 2010; Cull and Xu, 2003; Firth et al., 2009). In this section, I explore local favouritism towards capital-intensive industries in the allocation of credit by looking into the effect of tenure on firms' debt ratios across industries.

6.1 Profit Gap Rate

Cai and Liu (2009), using the same dataset as this paper, propose using the gap between a firm's imputed profits based on the *National Accounting Principle* (NAP) and its reported accounting profits based on the *General Accepted Accounting Principles* (GAAP) to measure the degree of tax evasion (or tax avoidance). $Pro_{i,t}$, the reported profit for firm i in year t, can be directly taken from the Annual Survey of Industrial Production. $ImpPro_{i,t}$, the imputed profit for firm i in year t, is calculated in the following way:

$$ImpPro_{i,t} = Y_{i,t} - Intermed_{i,t} - FC_{i,t} - Wage_{i,t} - CurDep_{i,t} - VAT_{i,t} \tag{6.1}$$

where $Y_{i,t}$ is the firm's gross output; $Intermed_{i,t}$ is intermediate inputs; $FC_{i,t}$ is financial charges; $Wage_{i,t}$ is the firm's total wage bill; $CurDep_{i,t}$ is current depreciation; and $VAT_{i,t}$ is value-added tax payments.

The profit gap $ProGap_{i,t} = ImpPro_{i,t} - Pro_{i,t}$. To be in line with the measurement of the effective VAT rate, I divide it by value-added to obtain the profit gap rate.

[21] The tenure effect on the change of VAT rate in Figure 5.4 is monotonic and seems different from the cyclical pattern as the tenure effect of the effect VAT rate in Figure 5.2. The difference between the two needs further exploration.

Conceptually, the profit gap can be decomposed into four components:

$$\begin{aligned} ProGap_{i,t} = & \, legitimate\ gap_{i,t} + tax\ evasion_{i,t} \\ & + earning\ management_{i,t} + accounting\ error_{i,,t} \end{aligned} \tag{6.2}$$

On the right-hand side of expression (6.2), there are three components in addition to *tax evasion*. The first term, *legitimate gap*, is the legal difference between NAP and GAAP.[22] The *accounting error* is assumed to follow a normal distribution with a zero mean. The *earning management* is not a big concern for our study as most firms in the sample are not publicly listed and do not have strong incentives to over-report earnings to deceive shareholders.[23] Given Expression (6.2), therefore, we only need to assume that the legitimate gap is not correlated with the tenure of the prefectural secretary of the CPC. Under this assumption, Expression (6.2) allows us to capture some fraction, if not all, of the misreporting of firms' profits.[24]

Figure 6.1: Distribution of Profit-Gap Rate

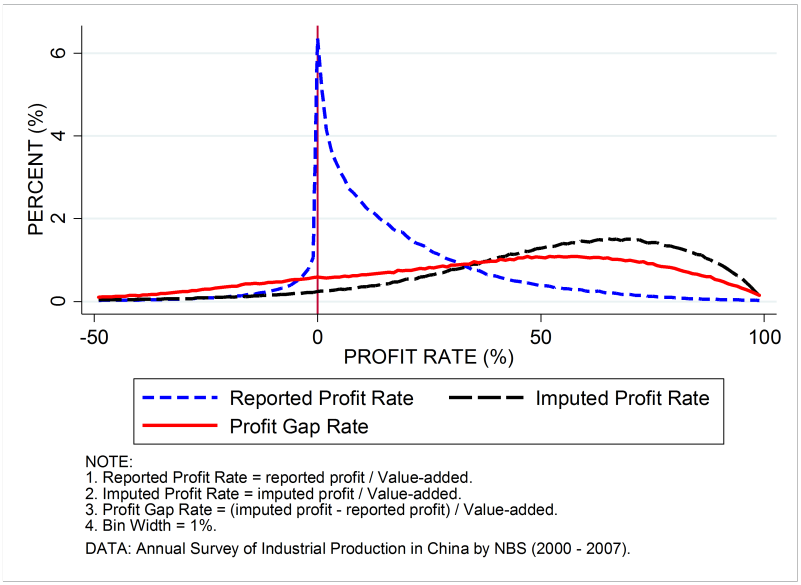


Figure 6.1 displays the distribution of imputed profit rate, reported profit rate, and the profit gap rate. These three rates are all normalized by firms' value-added. It should be noted that the reported profit rate has a significant spike around zero and is skewed

[22]The GAAP allows for more deductible items in calculating firms' accounting profits than the NAP does. Therefore the accounting profit $Pro_{i,t}$ is generally smaller than the imputed profit $ImpPro_{i,t}$. The *legitimate gap* between the two is approximately equal to: Manufacturing Expenses + Business Taxes and Surcharges + Operating Expense + Management Fees + Asset Impairment Loss + Loss from Changes in Fair Value + Investment Loss + Non-Business Expenditure - Non-Business Income - VAT.

[23]Studies related to this, such as Desai (2003, 2005), are mainly in the literature of corporate finance and accounting.

[24]Of course, Expression (6.2) cannot capture tax evasion that simultaneously changes the reported profits and imputed profits by the same magnitude.

to the right. In contrast, the distribution of the imputed profit rate is quite smooth. The spike suggests that some firms probably report very low but positive profits to evade tax.

Figure 6.2: Effective VAT Rate Vs. Profit-Gap Rate

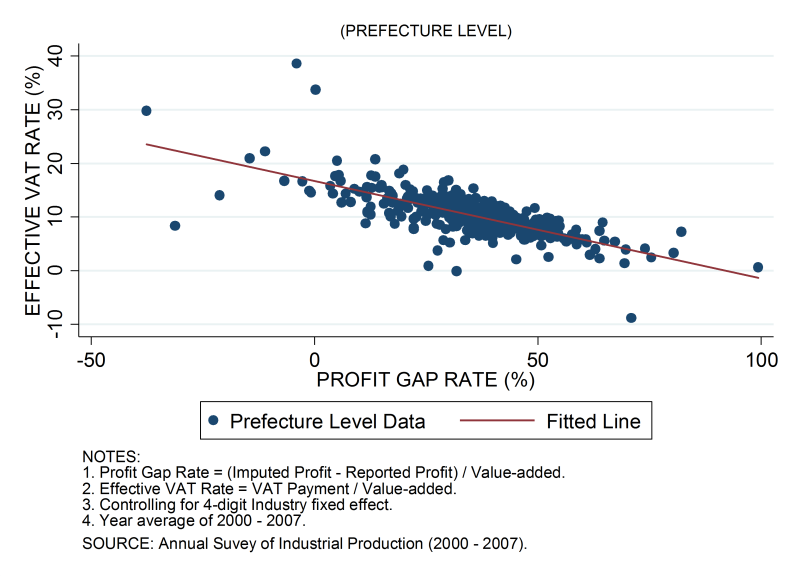


Figure 6.3: Tenure and Variation of Profit Gap Rate with Capital Intensity

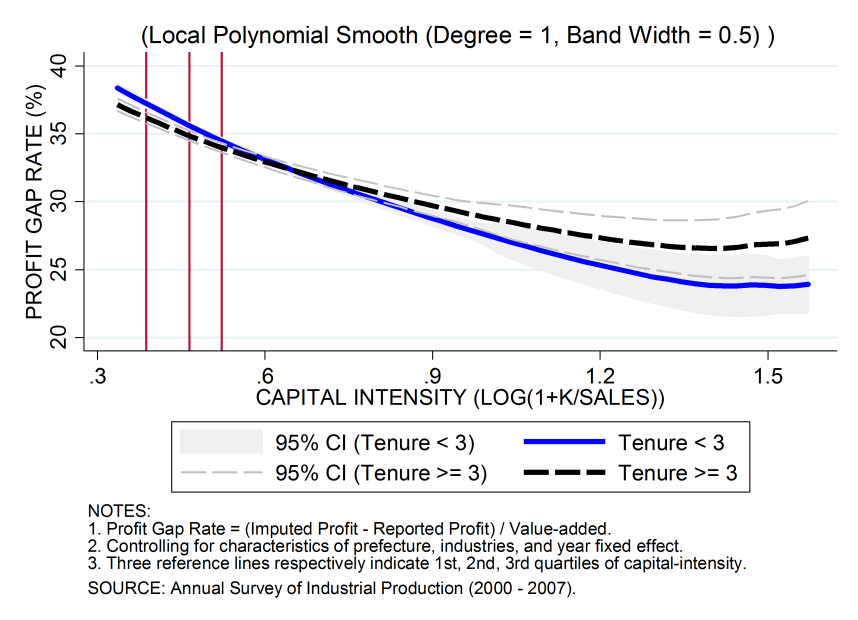


Figure 6.2 cross-validates the effective VAT rate and the profit gap rate at the prefecture level after controlling for the 4-digit industry fixed effect. We can see that the two rates are significantly and negatively correlated, implying that tax enforcement may consistently affect both the effective VAT rate and profit misreporting across regions. Since the effective VAT rate and the profit gap rate both have the value-added as their

denominators, the negative correlation between the two suggests that the variation of both variables is driven mainly by the numerators, which are potentially related to tax evasion, rather than being driven by the denominator.

As a first-pass test, following the same approach as in Figure 5.1, I provide in Figure 6.3 a visual representation of the relationship between the profit gap rate and capital intensity changes over tenure. The figure clearly suggests that the profit gap rate of capital-intensive firms goes up as the tenure rises, while the effective VAT rate of labour-intensive firms goes down. Similar to Figure 5.1, the figure shows that fewer than a quarter of the industries experience more lax tax treatment, while the remaining three-quarters face tougher treatment.

Table 4: Tenure Effect on Profit Gap Rate (INSERT)

Table 4 reports more formal empirical results.

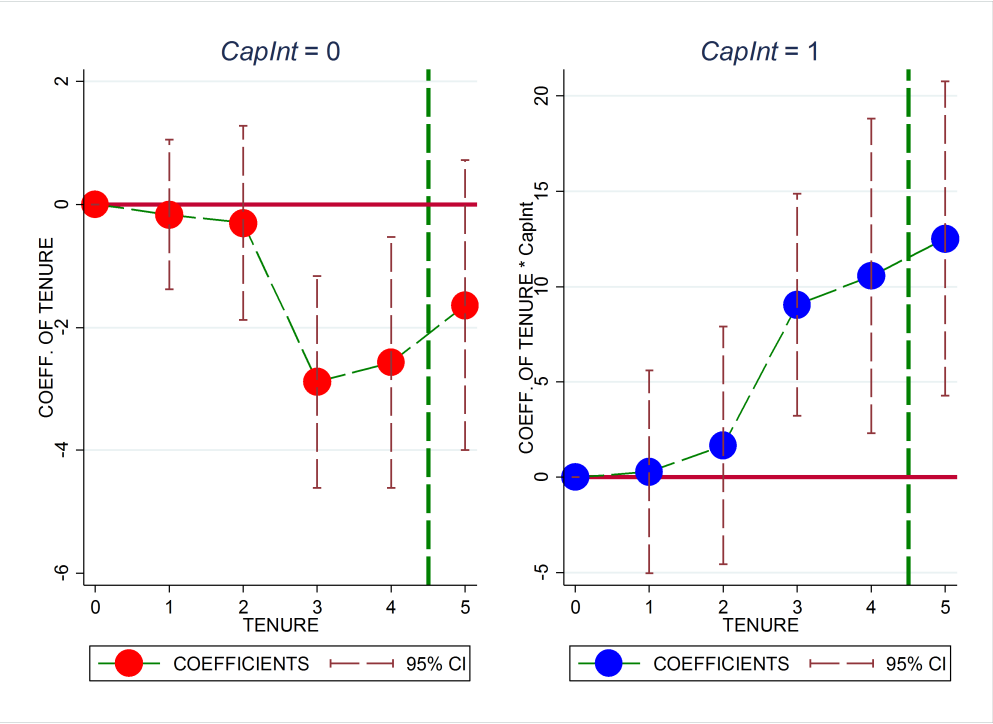
The results in columns (2) - (6) show pretty significant and robust tenure effects, which suggest that as the tenure rises, the profit misreporting rate increases for capital-intensive industries and decreases for labour-intensive ones. The results are consistent with the tenure effects on the effective VAT rate.

Two points should be noted for Panel B. First, unlike in the case of the effective VAT rate, some of the other industry characteristics appear to show a tenure effect. In particular, as tenure rises, an industry tends to report less profits if it has fewer state-owned firms, higher profits, bigger size, and a higher ratio of intermediate inputs relative to outputs. Second, the magnitude of the coefficient on "*Tenure*" rises significantly once the interaction of these characteristics with tenure is included.

Quantitatively, the results in Column (6) suggest that, when the prefectural secretary stays in office for additional year, the profit gap rate drops by 2.3% (or 9.2 percentage points from 40% to 30.8%) for labour-intensive industry with $CapInt = 0$, and rises by 9.7% (or 2.4 percentage points from 25% to 27.4%) for capital-intensive industry with $CapInt = 1$.

Figure 6.4 displays the dynamic effects of tenure on the profit gap rate for low and high capital intensity. The left picture shows the tenure effect on labour-intensive industries ("$CapInt_i = 0$"). The right picture shows the tenure effect on capital-intensive industries ("$CapInt_i = 1$"). The features in Figure 6.4 are quite consistent with those of the effective VAT rate in Figure 5.2.

Figure 6.4: Tenure Effect on Variation of Profit Gap Rate with Capital Intensity



6.2 Tenure Effect on Debt Ratio

The debt ratio is defined as a firm's total debt divided by total assets. It reflects a firm's willingness and ability to borrow from banks or other financial institutions.

Figure 6.5 shows how the relationship between the debt ratio and capital intensity changes over tenure. This time, the tenure effects do not show opposing directions depending on capital intensity. The picture suggests that the debt ratio of all industries goes up as the tenure rises, but more considerably and significantly for capital-intensive industries.

Table 5, with the same structure as Table 2, reports the tenure effects on firms' debt ratios based on Equation (5.1). Except for the coefficient on "*Tenure * CapInt*" in column (6), all other results in columns (2) - (6) show pretty significant and robust tenure effects, which again confirm the favouritism towards capital-intensive industries. As the tenure rises, the debt ratio rises for capital-intensive industries and falls for labour-intensive ones.

The results in Column (6) suggest that, when the prefectural secretary stays in office for additional year, the debt ratio drops by 6% (or 3.2 percentage points from 55.5% to 52.3%) for labour-intensive industry with $CapInt = 0$, and rises by 1% (or 0.5 percentage point from 55% to 55.5%) for capital-intensive industry with $CapInt = 1$.

It should be noted that there are some significant results on "*Tenure * other industry characteristics*", including ownership, profitability, LOG(V-added), and LOG(Sales). This means that, as tenure rises, industries with more non-state-owned and more profitable

firms tend to accumulate more debt.

Figure 6.5: Tenure Effect on Debt Ratio over capital intensity (LOWESS Smooth)

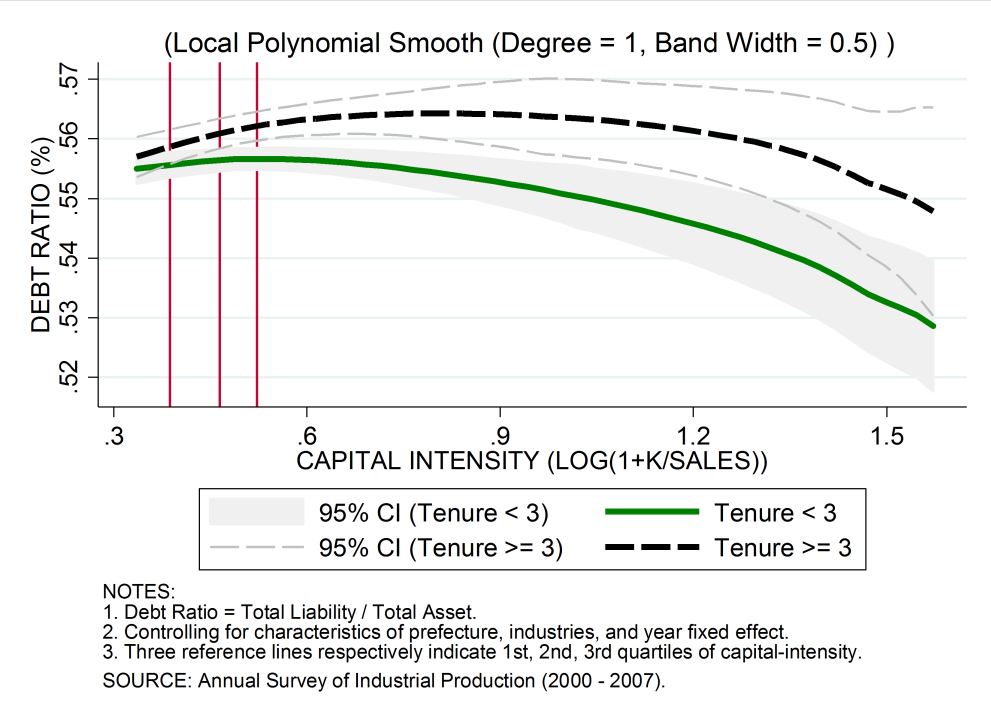


Table 5: Tenure Effect on Debt Ratio (INSERT)

Figure 6.6: Tenure Effect on Variation of Debt Ratio with Capital Intensity

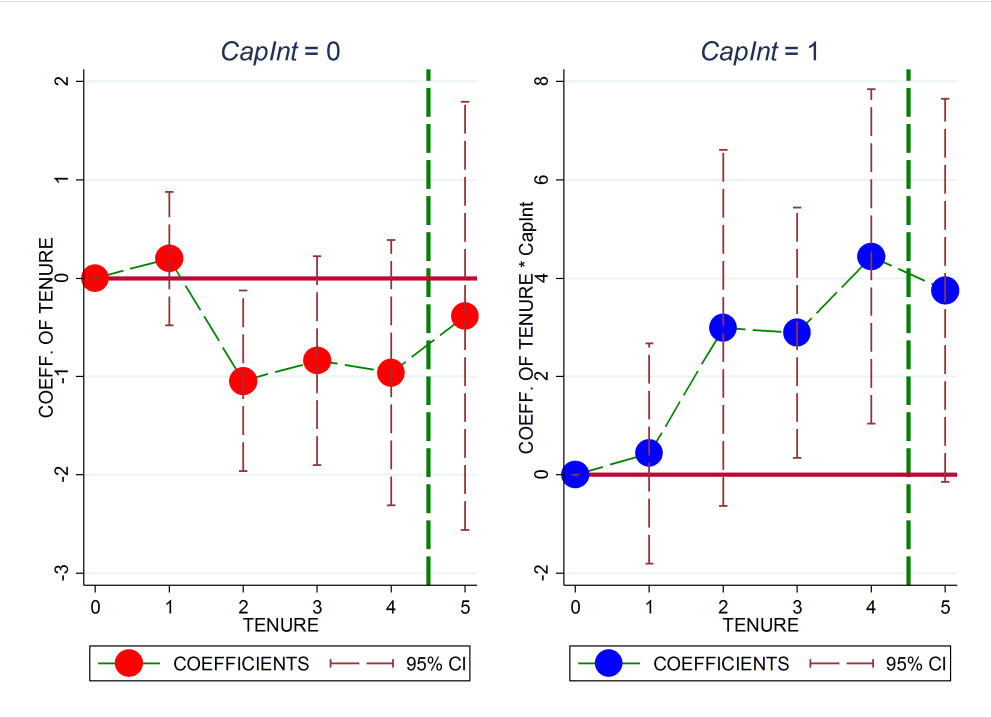


Figure 6.6 displays the dynamic tenure effects on the relationship between the debt ratio and capital intensity based on the dynamic version of the regression in Column (5) of Table 5. The left panel shows the tenure effect on labour-intensive industries ("$CapInt_i = 0$") and the right panel is for capital-intensive industries ("$CapInt_i = 1$"). From the graphs we can see that in year three the debt ratio falls for labour-intensive industries and rises for capital-intensive industries, and in year five it reverts to the original level for both industries.

7 Explaining the Tenure Effect

All of the results in Section 4 and Section 5 suggest that governments' preferential policies, both in taxation and credit rationing, shift in the direction of favouring capital-intensive industries as the tenure of a prefectural secretary increases.

What mechanism underlies the tenure effect? In this section, I explore several possibilities: corruption, politicians' career concerns and industrial policy, political connections, learning effects, and manipulation of value-added data. Although I cannot entirely rule out other possibilities, the evidence seems to be more in line with the explanation of corruption.

To help understand the mechanism of a tenure effect through corruption, I propose a multi-sector optimal taxation model with tax collection costs in a corrupt environment in Section 2. The model simply aims to explain the empirical findings in the paper.

7.1 Corruption

According to the Corruption Perception Index released by Transparency International, a global corruption watchdog, China ranked 80th in "Cleanness" among 177 countries worldwide in 2013.[25] But to measure corruption is a tough issue (Olken and Pande, 2012). And we know that, even in the same corrupt environment, the intensity of corruption varies considerably across firms (Svensson, 2003). Firms' incentives to participate in corruption could be determined by their various characteristics, such as capital intensity, ownership status, size, profitability, mobility, age, and so on. We therefore need to take these characteristics into consideration when studying corruption.

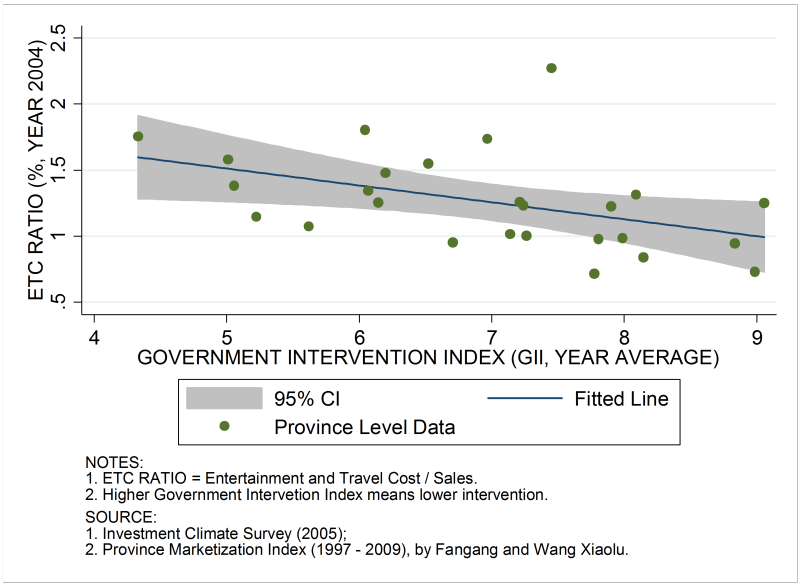
Following Cai et al. (2011), I use the ratio of a firm's entertainment and travel costs (ETC) to their total sales volume as a relevant indicator for its participation in corruption. In China, companies' entertainment costs includes the following four types of expenditure: (1) banquets and working meals; (2) souvenirs; (3) tickets, travelling cost, and other fees to tourist sites; and (4) expenses for business trips. Chinese firms spend extravagantly on entertainment to strengthen their relationship with government or with business partners.

[25] http://www.transparency.org/cpi2013/results.

It is well known that many bribes, pecuniary or non-pecuniary, are paid and reimbursed under the name of ETC expenditure in accounting practice.[26] With the data of the Investment Climate Survey in China, Cai et al. (2011) find that ETC is a mix that includes grease money to obtain better government services, or protection money to lower tax rates. It is therefore a plausible measure of corruption.

To provide further validation of ETC as an appropriate proxy for corruption, Figure 7.1 plots ETC against the Government Intervention Index (GII) at the province level.[27] A higher GII means a lower degree of government intervention and better market support. Theoretically, a province with lower GII has more rent-seeking opportunities and should be expected to have more expenditure on corruption, and we do see such a relationship in Figure 7.1.

Figure 7.1: ETC Vs. Government Intervention



If corruption is an underlying mechanism behind the favourable effect of tenure for capital-intensive firms, then the intensity of corruption of capital-intensive industries

[26]The media has reported that the entertainment costs of 2,469 listed Chinese companies in 2012 were 13.798 billion RMB yuan (2.26 billion US dollars), at an average of 0.9 million US dollars per company. Among these, the biggest ETC spender is China Life Insurance Company, spending 1.4 billion RMB (230 Million US dollars). In second place is China Railway Construction Corporation (CRCC), spending 837 Million RMB (137 million US dollars, or about 11% of its net profits). On August 1st, 2013, four judges of the Shanghai Higher Court, together with top leaders of Shanghai Construction Group Co.. Ltd, were caught calling prostitutes in a nightclub. It is alleged that the fees were to have been paid by Shanghai Construction Group, whose entertainment expenditure in 2012 was 178 million RMB yuan (29 nillion US dollars). (*China Youth Daily*, Aug 6th, 2013).

[27]The Government Intervention Index (GII), compiled by Fan and Wang (2011), is widely used in the Chinese economics literature. It is constructed along the following four sub-indices: 1. the role of market in resource allocation; 2. the burden of taxes and fees on farmers; 3. the government intervention to firms; 4. the burden on firms in addition to tax; 5. the size of government. The greater the GII, the lower the intensity of government intervention.

should rise faster over the CPC secretary's tenure than that of labour-intensive industries. To investigate this, I use the cross-sectional firm data from the Investment Climate Survey in China (2005).

Similar to Equation (5.1), the regression is specified as below to capture the cross-industry heterogeneous tenure effect on corruption:

$$\begin{aligned} y_{p,i} = \alpha \ + \ \eta_{r,s} \ + \ \lambda \cdot Tenure_p \ + \ \varphi \cdot Tenure_p \cdot CapInt_i + \\ \beta \cdot Tenure_p \cdot X_{p,i} \ + \ \rho \cdot CapInt_i \ + \ \psi \cdot Age\_Dummy_{p,t} + \\ \theta \cdot X_{p,i} \sum_{d=1}^{3} \omega_d \cdot Duration\_Dummy_{p,t}^{d} \ + \ \epsilon_{p,i} \end{aligned} \tag{7.1}$$

where the outcome variable $y_{p,i}$ is the ETC ratio. Subscripts f, p, r, i, s represent firm, prefecture, province, 2-digit-industry, 1-digit industry, respectively. $Tenure_p$ is the tenure of the secretary of the CPC in prefecture p in year 2004. $\eta_{r,s}$ captures the province-(1-digit)industry fixed effect. $CapInt_i$ is the median of capital intensity of 2-digit industry i in year 2004, with $CapInt_i$ normalized to zero for the least capital-intensive industries. $X_{p,i}$ is a vector of other firm characteristics of 2-digit industry i in prefecture p in year 2004, including ownership, profitability, firm age, logarithm of sales volume and logarithm of value-added. Reverse causality is a concern. For example, a corrupt firm may get more benefit and therefore become bigger and profitable. To deal with this problem, all these firm characteristics, except ownership, take the median of 2-digit industries of prefecture p in year 2004 (ownership takes the mean). Parameters of interests are λ and φ. Robustness standard errors are clustered at the province level. As in Equation (5.1), $\sum_{d=1}^{3} \omega_d \cdot Duration\_Dummy_{p,t}^{d}$ is introduced to avoid the selection problem.

Three differences should be noted between Equation (7.1) and Equation (5.1). First, due to data limitations, Equation (7.1) is just a cross-sectional regression. Second, $CapInt_i^2$ is not dropped in Equation (7.1). Even if it were included, its coefficient would not be significant and the main results would not be affected. Third, $Age\_Dummy_{p,t}$ is included in Equation (7.1). It is a dummy variable indicating whether the age of a prefectural secretary is greater than 50. I take 50 as the threshold not only because it is the median in the sample, but also because this appears to be the age after which prefectural secretaries are very unlikely to get promoted, and therefore they tend to be more corrupt.[28]

Table 6 shows the results, based on Equation (7.1), of the tenure effect of the prefectural secretary of the CPC on firms' ETC ratios at the firm level. The regression uses the cross-prefectural variation in secretary tenure. The results show that, in prefectures where the tenure is longer, the ETC ratio is significantly higher for capital-intensive industries. The tenure effect for labour-intensive industries in negative but is insignificant. Results in columns (2) - (5) shows that the relationship between the tenure effect and

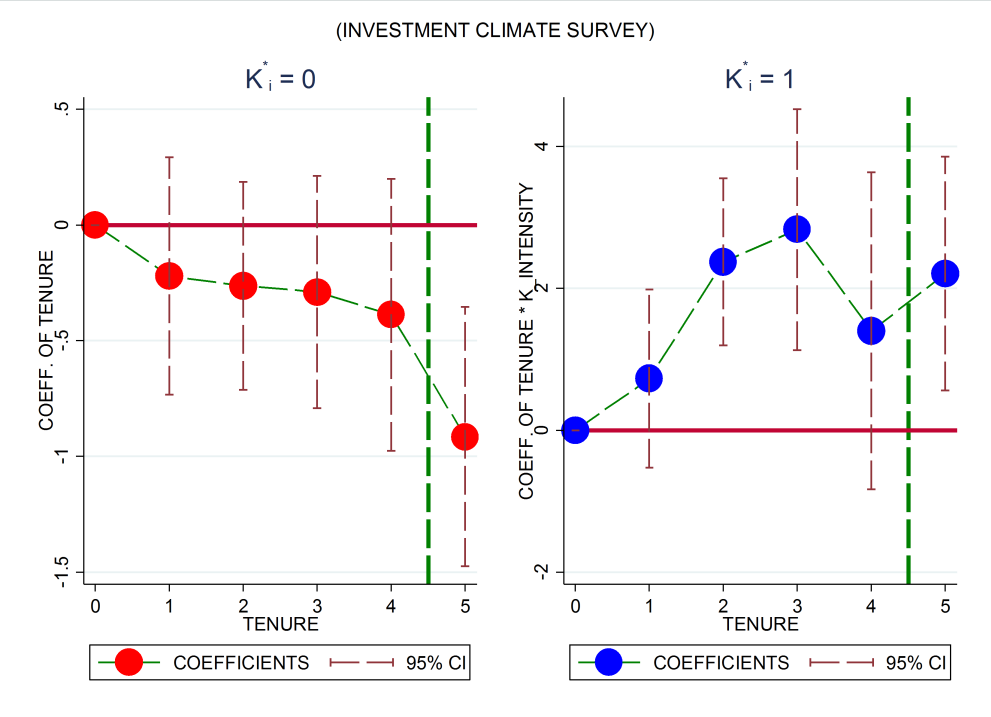
[28] In China, it is known as "50-year old phenomena" of prefectural secretaries.

capital intensity is robust to the inclusion of other firm characteristics. The results in Column (5) imply that, when the prefectural secretary stays in office for additional year, the ETC expenditure rises by 0.8% (or 0.09 percentage point from 10.9% to 11.8%) for capital-intensive industry with $CapInt = 1$.

Table 6: Tenure Effect on ETC (INSERT)

Figure 7.2 displays the dynamic tenure effect on ETC based on regression (7.1), with $Tenure_{p,t}$ replaced by a vector of dummy variables for each tenure year. The left panel shows the tenure effect on labour-intensive industries ("$CapInt_i = 0$") and the right panel is for capital-intensive industries ("$CapInt_i = 1$"). The graph shows that the ETC of capital-intensive industries goes up significantly in year three and remains higher until it falls in year five. In contrast, the ETC of labour-intensive industries remains unchanged until year five, when it falls sharply.

Figure 7.2: Tenure Dynamic Effect on ETC



Suppose that it is corruption that induces the favouritism towards capital intensity over the tenure, the question then is why capital-intensive firms have more incentives to participate in corruption. One possible way to answer this question is to look at the benefits associated with corruption at the firm level. These benefits include a VAT reduction and access to credit.

Based on the cross-sectional firm data from the Investment Climate Survey in China

(2005), the baseline regression can be specified as follows.

$$y_f = \alpha + \eta_{p,s} + \lambda \cdot ETC_f + \varphi \cdot ETC_f \cdot K_f + \beta \cdot ETC_f \cdot X_f + \rho \cdot K_f + \theta \cdot X_f + \epsilon_f \tag{7.2}$$

where the notations are the same as those in Equation (5.1). The outcome variables y_f are "Effective VAT Rate" or "Collateral dummy", respectively. The latter indicates whether the collateral is required for an overdraft or a bank loan. $\eta_{p,s}$ is the prefecture-(1-digit)industry fixed effect. X_f and its interaction with ETC_f are included as controls. It should be noted that all the firm characteristics used in the regression are at the firm level and are observable to government officials, because we want to know the firm-specific return to ETC and manager's political status. It should be expected that the return to corruption and political connections is higher for capital-intensive firms.

Table 7 and Table 8 report two types of observable preferential policies – VAT reduction and credit access, respectively – that capital-intensive firms with higher ETC can obtain.

Table 7 displays how the association between the effective VAT rate and ETC expenditure varies with capital intensity. Both the baseline regression results in Panel A (based on equation (7.2)) and the results in Panel B show the robustly negative coefficient on "$ETC * CapInt$". It suggests, that for the same amount of ETC expenditure, capital-intensive firms can obtain greater VAT reductions than labour-intensive firms. The results in Panel B also show that other industrial characteristics are not channels through which the ETC affects the effective VAT rate. Quantitatively, the results in Column (8) suggest that, when the ETC ratio rises by 1 percentage point, then the effective VAT would drop by 0.0153 percentage point (or around 14%) for capital-intensive industry with $CapInt = 1$.

Table 8 reports how the relation between collateral requirements and ETC expenditure varies with capital intensity. Both the baseline regression results in Panel A (based on equation (7.2)) and the results in Panel B show the robustly negative coefficient on "$ETC * CapInt$". It also suggests that other characteristics seem not to be the channels through which the ETC affects access to finance. The results in Column (8) imply that, when the ETC ratio rises by 1 percentage point, then the probability of collateral requirement for bank loan would drop by 0.71 percentage point for capital-intensive industry with $CapInt = 1$.

Table 7: Return to ETC – Effective VAT Rate (INSERT)

Table 8: Return to ETC – Collateral Requirement for Bank Loan (INSERT)

One caveat should be noted. The results in Table 7 and Table 8 confirm that capital-

intensive firms obtain more returns for the same ETC expenditure. But it seems puzzling that the labour-intensive firms' ETC has a negative return in some regressions, for both the effective VAT rate (columns (1) - (6) of Table 7) and collateral requirement (columns (1), (3) - (5) of Table 8). However, it is not significant in column (8) once all characteristics are controlled for. Although this is not a big concern as this paper is more interested in the heterogeneity with capital intensity – that is, the coefficient of "$ETC*CapInt$" rather than the coefficients of "ETC" – I would still like to provide two possible explanations for this result. One possibility is the general equilibrium effect of corruption. Given the fact that the total rent-seeking opportunities are fixed, it would be a zero-sum game if both capital-intensive firms and labour-intensive were to compete for the preferential treatment from the local government. If the capital-intensive firms spend more on ETC and win the game, then it would be the labour-intensive firms who would have to pay the higher tax, even if they have also already paid the ETC of corruption. Another possible explanation is that labour-intensive firms can actually get positive returns to ETC in other aspects that are not observed by statisticians, even though the returns in terms of a VAT reduction and an ease in credit appear to be negative.

7.2 Other Channels

In addition to heterogeneous corruption depending on capital intensity, there are a couple of potential mechanisms underlying the tenure effect. I discuss these possibilities in this section. However, the evidence suggests that most of these mechanisms are unlikely to be functioning.

7.2.1 Politicians Career Concern and Industrial Policy

Several authors argue that local politicians in China compete in GDP growth in order to achieve promotion (Li and Zhou, 2005; Chen, Li et al. 2005). For politicians, one way to boost short-run GDP growth is to stimulate investment in physical assets, and capital-intensive industries provide good leverage to do so. To verify this explanation, I studied the tenure effect on firms' long-term investments based on Equation (5.1). However, the regression results show no significant effect.

Another reason for the preferential policy towards capital-intensive firms could be local governments "picking the winner" and wanting to divert resources to the successful capital-intensive firms. However, it is difficult to reconcile this explanation with the empirical results. I studied the distribution of firms' profitability (measured by returns-on-assets, ROA) and productivity (measured by total factor productivity of revenue, TFPR) according to capital intensity.[29] The results show that both ROA and TFPR of capital-intensive

[29]Following Hsieh and Klenow (2009), the total factor productivity of Revenue (TFPR) of firm i is

firms are lower than for labour-intensive firms. And the regression of the tenure effect based on Equation (5.1) suggests that capital-intensive firms become even less profitable (lower ROA) and less productive (lower TFPR) as the tenure of the prefectural secretary rises. If the local politicians are picking the winner, why are preferential policies given to unsuccessful capital-intensive firms? And why do these firms become even more unsuccessful after obtaining the preferential policies?

7.2.2 Political Connections

Many studies use the political background of firms' top leaders to measure political connections, and study the benefits to firms (Faccion, 2006; Khwaja and Mian, 2005). In China, it is not only the firm that has an incentive to hire a top leader with a political background; the government would also like to appoint someone – either a former government official or a professional manager – as the top leader of a firm in order to tighten its grip on the firm's operations. To capture these political connections and government intervention, I also use the status of the firm's general manager – a dummy variable indicating whether or not the manager is appointed by the government – which is available from the Investment Climate Survey (2005) in China conducted by the World Bank. The likelihood of appointment of a politically connected manager does not change over the tenure of the prefectural secretary, in respective of a firm's capital intensity.

7.2.3 Learning Effect

One may argue that the shifts in taxation from capital-intensive firms to labour-intensive firms as the tenure rises could be due to the secretary, or some lower level officials newly appointed by the secretary, accumulating skills in taxation during their leadership. However, several hard facts in China go against this argument. First, prefectural secretaries rarely need to learn how to tax; it is the tax inspectors who collect tax from firms directly. Also, most grassroots taxation staff are very stable in their positions. The impact of prefectural secretaries' turnover on their position and taxation skill is very limited, if not inexistent. Second, lacking the detailed information and professional skill in taxation, the prefectural secretary in practice would only set the annual tax revenue target for the local office of the SAT, rather than painstakingly lecturing tax inspectors on how to tax differently across industries. Third, if the shift in taxation from capital-intensive firms to labour-intensive firms is due to learning effects, it must require the prefectural secretary to learn how to tax labour-intensive firms more quickly than capital-intensive firms. This is at odds with the common understanding that it is easier to learn how to tax capital-intensive firms than labour-intensive firms (Gordon and Li, 2009).

calculated by $TFPR_i = \frac{VA_i}{K_i^{\alpha} \cdot L_i^{1-\alpha}}$, where α is the industry-level capital share.

7.2.4 Manipulation of Value-added Data

Figure 6.2 has already shown that the cross-regional variation of the effective VAT rate and the profit gap rate are mainly driven by the numerator (which is related to tax enforcement), rather than the denominator (firms' reported value-added). However, one may still suspect that the tenure effect may come from the manipulation of firms' reported value-added rather than from tax enforcement. Even though it is true that local politicians in China may require firms to over-report their value-added in order to fabricate local economic performance and achieve promotion, this explanation is inconsistent with another two facts. First, if the manipulation of firms' value-added is the cause of the tenure effect on the effective VAT rate, then the opposing tenure effects for capital-intensive industries and labour-intensive industries must imply that the value-added is rigged upwards for some industries and rigged downwards for others. Obviously, it would be extremely unlikely for any local politician to have done this. Second, if the tenure effect comes from the variation of value-added (the denominator of both the effective VAT rate and the profit gap rate), then the tenure effect should display the same direction for both the effective VAT rate and the profit gap rate. The results, however, show opposing directions for these two rates.

7.2.5 Input-output Structure and Export Refund

Exports can earn a firm VAT refunds. In addition to this, both the analysis in Section 4.1.1 and the regression results show that a higher input-output ratio for an industry may also reduce its effective VAT rate. Does the tenure effect on the effective VAT rate work not through tax enforcement, but instead through input-output structure or export refunds? To examine whether this is possible, I study the tenure effect on input-output structure and the export ratio with Equation 5.1. The dependent variable is replaced with either the input-output ratio or the export ratio at the county-(2-digit)industry-year level. The results show no significant tenure effect on export refunds. This implies that export refunds are not an alternative channel for the tenure effect. Although the tenure effect is significant for capital-intensive industries in some regressions at the 10% level, the coefficient φ is negative. Because a lower input-output ratio leads to a higher effective VAT rate, if the input-output channel does exist, it would actually attenuate the tenure effect and bias the tenure effect towards zero. Therefore, the true tenure effect would be even stronger than I have estimated.

8 Conclusion

This paper confirms that tax enforcement on firms and credit allocation over them vary with local political cycles, as represented by the tenure of the prefectural secretary

of the Communist Party, who is the *de facto* local governor in China.

The paper begins with the cross-industry distribution of the effective VAT rate. VAT is a technologically sound tax that is in operation across around 150 countries. China adopted VAT in 1994 and has established a nationwide VAT special invoice cross-checking system. To prevent local intervention, VAT administration has always been under the charge of the central government-led State Administration of Taxation (SAT). However, VAT enforcement in China is still proving to be problematic. The empirical results show rising favouritism towards capital-intensive industries over the tenure of the prefectural secretary. That is, as the tenure rises, the effective VAT rate falls for capital-intensive industries and rises for labour-intensive industries. In addition, the paper uses the nationwide abolition of agricultural tax in 2005 as a natural experiment in which local governments were forced to raise the effective VAT rate for local manufacturing firms. The results show that the VAT rate rises less for capital-intensive industries and more for labour-intensive ones the longer the tenure of the prefectural secretary.

The favouritism towards capital-intensive industries appears to extend to two additional outcomes. First, the tenure effect on profit misreporting suggests that capital-intensive firms are more likely to under-report their profits to evade corporate income tax the longer the tenure of the prefectural secretary, and vice versa for labour-intensive firms. Second, the tenure effect on credit allocation shows that credit is directed from labour-intensive firms to capital-intensive ones as the tenure rises.

The paper finally points out that corruption is a possible underlying mechanism of the tenure effect on tax enforcement and credit allocation. Other channels, such as prefectural secretaries' career concerns, industrial policy, political connections or learning effects, seem inconsistent with some of the facts and evidence. Of course, given the indirect measurement of corruption and the complexity of political institutions, this paper is unable to provide ample evidence in this regard. More work needs to be done in the future once better data are available and better identification is possible.

The paper presents a broader and more complete picture of the general equilibrium effect (or external effect) of corruption. It examines the impact of corruption not only on corrupt firms but also on incorrupt (or less corrupt) ones, while most of the existing literature only study the direct benefits to the corrupt firms. The results imply, given the total tax revenue target that the tax inspectors have to achieve, that more lax tax enforcement for capital-intensive firms must require tougher enforcement for their labour-intensive counterparts. By the same token, given the credit constraints faced by local governments, access to finance has to shift from labour-intensive firms to capital-intensive firms. In a general sense, because politicians can control only limited resources, corruption creates negative externalities for incorrupt, or less corrupt, firms.

The findings in this paper may also enrich our understanding of the relationship be-

tween legal capacity and fiscal capacity, which is emphasized by Besley and Persson (2009, 2011, 2013). As the paper demonstrates, even though developing countries can copy tax codes from the developed world, this cannot be well enforced in a country where the rule of law is not respected by powerful politicians.

To sum up, the paper contributes to establishing the facts on the cross-industry distribution of VAT enforcement over local political cycles. However, due to the lack of data and a proper identification strategy, several related questions remain unanswered. For example, how is the local political network affected by prefectural secretary turnover? What is the incentive for local political intervention in VAT? Is it related to the age of secretaries and their chances of promotion? To what degree does local political intervention damage the nation's taxation capacity? Why is the return to ETC higher for capital-intensive industries? These questions are still calling for further research.

References

1. Adhikari, A., Derashid, C. and Zhang, H. (2006) "Public policy, political connections, and effective tax rates: Longitudinal evidence from Malaysia", *Journal of Accounting and Public Policy*, 25(5), pp. 574-595.

2. Aizenman, J. and Jinjarak, Y.. (2008) "The collection efficiency of the Value Added Tax: Theory and international evidence", *The Journal of International Trade and Economic Development*, Taylor and Francis Journals, 17(3), pp. 391-410.

3. Allen, F., Qian, J. and Qian, M. (2005) "Law, finance, and economic growth in China", *Journal of Financial Economics*, 77(1), pp. 57-116.

4. Almunia, M. and Lopez-Rodriguez, D. (2014) "Heterogeneous Responses to Effective Tax Enforcement: Evidence from Spanish Firms", *Working paper*, University of Warwick.

5. Andreoni, J., Erard, B. and Feinstein, J. (1998) "Tax Compliance", *Journal of Economic Literature,* 36(2), pp. 818-860.

6. Bertrand, M., Duflo, E. and Mullainathan, S. (2004) "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1), pp. 249-275.

7. Besley, T. and Persson, T. (2009) "The Origins of State Capacity: Property Rights, Taxation, and Politics", *American Economic Review*, 99(4), pp. 1218-44.

8. Besley, T. and Persson, T. (2011) "*Pillars of Prosperity: The Political Economics of Development Clusters,*" Princeton University Press.

9. Besley, T. and Persson, T. (2013) “Taxation and Development”, *Handbook of Public Economics* Vol. 5, Elsevier, Amsterdam. in Alan Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez (eds.), Elsevier, Amsterdam.

10. Cai, H. and Liu, Q. (2009) "Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms", *Economic Journal*, 119(537), pp. 764-795.

11. Cai, H., Fang, H. and Xu, C. L. (2011) "Eat, Drink, Firms, Government: An Investigation of Corruption from the Entertainment and Travel Costs of Chinese Firms", *Journal of Law and Economics*, University of Chicago Press, 54(1), pp. 55 - 78.

12. Chen, S. X. G. (2015a) “VAT Dispersion and Aggregate Efficiency in China”, *Working Paper*, London School of Economics, Accepted by 2014 Annual Conference of Royal Economic Society, UK.

13. Chen, S. X. G. (2015b) “The Effect of a Fiscal Squeeze on Tax Enforcement: Evidence from a Natural Experiment in China”, *Working Paper*, London School of Economics.

14. Chen, Y., Li, H. B. and Zhou, L. A. (2005) "Relative performance evaluation and the turnover of provincial leaders in China", *Economics Letters*, 88(3), pp. 421-425.

15. Cull, R. and Xu, C. L. X. (2003) "Who gets credit? The behavior of bureaucrats and state banks in allocating credit to Chinese state-owned enterprises", *Journal of Development Economics*, 71(2), pp. 533-559.

16. Desai, M. A. (2003) "The Divergence between Book and Tax Income", *Tax Policy and the Economy* 17: 169–206.

17. Desai, M. A. (2005) “The Degradation of Reported Corporate Profit”, *Journal of Economic Perspectives*, 19 (1), 171–192.

18. *Economist* (2014) “Our Cronyism Index – Planet Plutocrat”, March 15th - 21st.

19. Faccio, M. (2006) "Politically Connected Firms", *American Economic Review*, American Economic Association, vol. 96(1), pages 369-386, March.

20. Faccio, M. (2007) “The Characteristics of Politically Connected Firms”, *Working Paper*, Purdue University.

21. Fan, G. and Wang, X. L. (2011) “Province Marketization Index in China (1997 - 2009)”, *National Economic Research Institute*, China Reform Foundation.

22. Fan, J. P. H., Wong, T. J. and Zhang, T. Y. (2007) "Politically connected CEOs, corporate governance, and Post-IPO performance of China's newly partially privatized firms", *Journal of Financial Economics*, 84(2), pp. 330-357, May.

23. Firth, M., Lin, C., Liu, P. and Wong, S. (2009) "Inside the Black Box: Bank Credit Allocation in China's Private Sector", *Journal of Banking and Finance*, 33(6), pp. 1144-1155.

24. Fisman, R. (2001) "Estimating the Value of Political Connections", *American Economic Review*, 91(4), pp. 1095-1102.

25. Fisman, R. and Wei, S. J. (2004) "Tax Rates and Tax Evasion: Evidence from 'Missing Imports' in China", *Journal of Political Economy*, 112(2), pp. 471-496.

26. Fisman, R. and Svensson, J. (2007) "Are corruption and taxation really harmful to growth? Firm level evidence", *Journal of Development Economics*, 83(1), pp. 63-75.

27. Fisman, R. and Wang, Y. X. (2015) "The mortality cost of political connections", *Review of Economic Studies*, Forthcoming.

28. Fisman, R. and Wang, Y. X. (2014) "Corruption in Chinese Privatizations", *Journal of Law, Economics, and Organization*, pp. 1- 29.

29. Gordon, R. and Li, W. (2009) "Tax structures in developing countries: Many puzzles and a possible explanation", *Journal of Public Economics,* 93(7-8), pp. 855-866.

30. Keen, M. and Lockwood, B. (2010) "The value added tax: Its causes and consequences", *Journal of Development Economics*, 92(2), pp. 138-151.

31. Khwaja, A. I. and Mian, A. (2005). "Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market", *The Quarterly Journal of Economics*, 120(4), pp. 1371-1411.

32. Khwaja, A. I. and Mian, A.(2011) "Rent Seeking and Corruption in Financial Markets", *Annual Review of Economics,* 3(1), pp. 579-600.

33. Khwaja, A. I., Olken, B. A. and Khan, A. (2014) "Property Tax Experiment in Punjab, Pakistan: Testing the Role of Wages, Incentives and Audit on Tax Inspectors' Behavior", *Working Paper.*

34. Kleven, H. J., Knudsen, M., Kreiner, C. T., Pedersen, S. and Saez, E. (2011) "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark", *Econometrica*, 79, pp. 651-692.

35. Li, H. B., Meng, L. S., Wang, Q. and Zhou, L. A. (2008) "Political connections, financing and firm performance: Evidence from Chinese private firms", *Journal of Development Economics*, 87(2), pp. 283-299.

36. Li, H. B. and Zhou, L. A. (2005) "Political turnover and economic performance: the incentive role of personnel control in China", *Journal of Public Economics*, 89(9-10), pp. 1743-1762.

37. OECD. (2014) "Foreign Bribery Report An Analysis of the Crime of Bribery of Foreign Public Officials", OECD Publishing. Paris: OECD.

38. Olken, B. A. and Pande, R. (2012) "Corruption in Developing Countries", *Annual Review of Economics*, 4(1), pp. 479-509.

39. Poncet, S., Steingress, W. and Vandenbussche, H. (2010) "Financial constraints in China: Firm-level evidence", *China Economic Review*, 21(3), pp. 411-422.

40. Pomeranz, D. (2013) "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax", *NBER Working Papers* 19199, National Bureau of Economic Research, Inc.

41. Koren, M., Szeidl, A. and Vedres, B. (2014) "Political Connections and Procurement in Hungary", Central European University, *Working paper*.

42. Slemrod, Joel., and Yitzhaki, S. (2002) "Tax Avoidance, Evasion, and Administration", In *Handbook of Public Economics*, Volume 3, edited by A.J. Auerbach and M. Feldstein, 1423-1470, North Holland.

43. Slemrod, J. (2007) "Cheating Ourselves: The Economics of Tax Evasion", *Journal of Economic Perspectives*, 21(1), pp. 25-48.

44. Slemrod, J., Blumenthal, M. and Christian, C. W. (2001) "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota", *Journal of Public Economics*, 79(3), 455-483.

45. Svensson, J. (2003) "Who Must Pay Bribes and How Much? Evidence from a Cross-Section of Firms," *The Quarterly Journal of Economics*, MIT Press, 118(1), pp. 207-230.

46. Wu, W. F., Wu, C. F., and Rui, M. (2009) "Government Background of Top Leaders in China's Listed Companies and Tax Preferential Policy", *Management World*, Issue 3. (In Chinese)

47. Xu, C. G. (2011) "The Fundamental Institutions of China's Reforms and Development," *The Journal of Economic Literature*, 49(4), pp. 1076-1151.

Table 1: Summary Statistics

Panel A: Annual Survey of Industrial Production in China (2000 - 2007)

	Sample Size	Mean	St. Dev.	Min	Max	p10	p50	p90
Effective VAT Rate[1]	139,287	10.34	9.66	-153.22	697.38	0.98	8.74	20.87
Profit Gap Rate[2]	139,287	34.67	42.85	-2783.55	2813.98	-12.3	40.19	75.92
Debt Ratio[3]	139,287	55.77	28.94	-120.02	1585.95	22.19	55.28	86.37
CapInt[4]	139,287	0.18	0.21	0	1.24	0.01	0.13	0.32
Ownership[5]	139,287	0.89	0.18	0	1	0.7	0.95	1
Profitability[6]	139,287	0.09	0.65	-16.8	176.27	0.03	0.06	0.14
Firm Age[7]	139,287	6.36	4.71	0	95	3	6	9
LOG(V-added)[8]	139,287	9.05	0.69	3.76	16.31	8.33	8.96	9.85
LOG(Sales)[9]	139,287	10.1	0.7	3.65	16.4	9.36	10.04	10.9
Input - output Ratio[10]	139,287	0.76	0.17	0	21.3	0.69	0.76	0.84
Export - Sales Ratio[11]	139,287	0.02	0.13	0	10	0	0	0

Panel B: Investment Climate Survey in China (2005)

	Sample Size	Mean	St. Dev.	Min	Max	p10	p50	p90
Effective VAT Rate[1]	1,271	9.19	45.02	-1487.88	351.19	2.06	8.70	17.06
ETC Ratio[12]	1,271	1.09	1.18	0	15.85	0.24	0.78	2.1
Govt-appointed Manager[13]	1,270	1.88	0.2	1	2	1.67	2	2
Collateral Dummy[14]	1,271	0.51	0.29	0	1	0	0.5	1
CapInt[4]	1,271	0.13	0.1	0	0.54	0.03	0.13	0.32
Ownership[5]	1,271	0.95	0.21	0	1	1	1	1
Profitability[6]	1,271	0.04	0.8	-27.73	1.33	0	0.05	0.18
Firm Age[7]	1,271	10.98	9.93	2	86	4	8	20
Manager Tenure[15]	1,271	5.55	2.64	1	27	3	5	9
LOG(V-added)[8]	1,270	10.01	1.61	4.62	17.49	8.2	9.83	11.96
LOG(Sales)[9]	1,271	10.98	1.55	5.71	17.51	9.19	10.89	13

Panel C: Political Turnover of Prefecture Secretary of CPC (2000 - 2010)

	Sample Size	Mean	St. Dev.	Min	Max	p10	p50	p90
Tenure[16]	2,800	1.91	1.76	0	14	0	2	4
Duration[17]	2,800	4.13	1.96	0	15	2	4	7
Age[18]	2,800	49.41	4.05	25.95	59.07	44.01	49.95	54.07

Panel D: County Public Finance Statistics (2000 - 2007)

	Sample Size	Mean	St. Dev.	Min	Max	p10	p50	p90
Agr[19]	19,950	0.25	0.23	-0.81	1.21	0.01	0.22	0.56
LOG(Pop)[20]	8,898	3.47	0.94	0	5.78	2.2	3.61	4.52
LOG(GDP/Pop)[21]	8,823	8.44	0.86	-2.73	12.4	7.59	8.43	9.41
Fiscal Pressure[22]	7,657	501.82	844.06	-15818.85	10399	-7.22	375.67	1193.59

Notes:

1. "Effective VAT Rate" is defined as "Payable VAT / value-added × 100".

2. "Profit Gap Rate" is measured by "profit gap / value-added × 100", where "profit gap" is the "imputed profit" minus the "reported profit". See Section 6.1 for more details.

3. "Debt Ratio" is measured by "total liability/ total assets × 100".

4. "CapInt" is normalized capital-intensity measured by $Log(1 + capital/sales)$, where $Capital$ is measured by "total asset" in Panel A, and by "net fixed asset / sales" in Panel B. The measurements are different due to lack of consistent data. By subtracting the minimum $Log(1+capital/sales)$ among all industries, the "CapInt" of the least capital-intensive industry is normalized to zero.

5. "Ownership" is dummy variable which is equal to 1 if the firm is not a state-owned enterprise.

6. "Profitability" is measured by "profit / value-added", indicating the degree of competition of 2-digit industry.

7. "Firm Age" is the years since the firm was set up.

8. "LOG(V-added)" is the logarithm of a firm's value-added.

9. "LOG(Sales)" is the logarithm of a firm's sales volume.

10. "Input-output Ratio" is defined as "intermediate inputs / total value of products and services".

11. "Export-sales Ratio" is defined as "export / sales volume".

12. "ETC" is defined as "entertainment and travel expenditure / total sales × 100" of the firm.

13. "Govt-appointed Manager" is a dummy taking value 1 if the firm's general manager is appointed by the government.

14. "Collateral" is a dummy variable taking value 1 if collateral is required for an overdraft or loan.

15. "Manager Tenure" is total years since firm's general manager was in office.

16. "Tenure" indicates the years for which the prefecture secretary has been in current office.

17. "Duration" is the total years of service of the prefectural secretary, indicating the length of periods from his/her entering to leaving between two consecutive political turnovers.

18. "Age" refers to the age of the prefectural secretary of CPC.

19. "Agr" is the net tax revenue loss due to the abolition of agricultural tax. Refer to Expression (5.2) for its definition.

20. "LOG(Pop)" is the logarithm of a county's total population.

21. "LOG(GDP/Pop)" is the logarithm of a county's GDP per capita.

22. "Fiscal Pressure" = (total expenditure − construction expenditure − net government transfer revenue − net non-tax revenue) / total population at the county government level of each year;

Table 2: Tenure Effect on Effective VAT Rate

	Dependent Variable: Effective VAT Rate					
	Panel A: Baseline			Panel B: Control for *Tenure* × *Other Characteristics*		
	(1)	(2)	(3)	(4)	(5)	(6)
Tenure	0.09	0.17	0.16	0.64	0.71	0.81
	(0.06)	(0.07)**	(0.07)**	(0.66)	(0.39)*	(0.40)**
Tenure	NO	-0.41	-0.36	-0.53	-0.29	-0.29
× CapInt		(0.15)***	(0.15)**	(0.17)***	(0.14)**	(0.15)*
	Industrial Trends					
CapInt × Year	NO	NO	-0.70	-0.70	-0.61	-0.58
			(0.13)***	(0.13)***	(0.11)***	(0.11)***
	Fixed Effects					
Province-Year FE	YES	YES	YES	YES	YES	YES
Prefecture-(1)Indu FE	YES	YES	YES	YES	YES	NO
County-(1)Indu FE	NO	NO	NO	NO	NO	YES
	Secretary's Controls					
Dummy (Duration < 5)	0.20	0.20	0.19	0.19	0.27	0.36
	(0.16)	(0.15)	(0.16)	(0.15)	(0.16)*	(0.17)**
Dummy (Duration > 5)	-0.06	-0.06	-0.07	-0.07	-0.19	-0.12
	(0.23)	(0.22)	(0.23)	-0.23	-0.2	-0.22
Tenure Square	-0.02	-0.02	-0.02	-0.02	0.00	0.00
	(0.01)*	(0.01)*	(0.01)*	(0.01)*	(0.01)	(0.01)
	Tenure × *Other Characteristics*					
	Not Included			*Included*		
Tenure				-0.28	-0.17	-0.29
× Ownership				(0.25)	(0.19)	(0.21)
Tenure				0.01	-0.03	-0.02
× Profitability				(0.07)	(0.04)	(0.05)
Tenure				0.00	0.00	-0.01
× Firm Age				(0.01)	(0.00)	(0.00)
Tenure				0.08	0.07	0.04
× LOG(Value-added)				(0.12)	(0.08)	(0.09)
Tenure				-0.08	-0.11	-0.08
× LOG(Sales)				(0.13)	(0.08)	(0.08)
Tenure				-0.24	-0.03	-0.02
× Input-Output Ratio				(0.32)	(0.11)	(0.12)
Tenure				-0.01	-0.11	-0.08
× Export Ratio				(0.15)	(0.12)	(0.13)

(Continued on next page)

Table 2: Tenure Effect on Effective VAT Rate (Continued)

	2-Digit Industry Characteristics					
	(1)	(2)	(3)	(4)	(5)	(6)
CapInt	12.27	12.99	1411.89	1416.12	1230.14	1175.81
	(1.42)***	(1.46)***	(259.80)***	(262.10)***	(217.68)***	(227.84)***
CapInt Square	-8.35	-8.35	-8.48	-8.5	-6.98	-7.12
	(1.51)***	(1.53)***	(1.52)***	(1.52)***	(1.24)***	(1.25)***
Ownership	-0.94	-0.95	-0.96	-0.46	-0.69	-0.45
	(0.66)	(0.66)	(0.66)	(0.85)	(0.70)	(0.72)
Profitability	-0.14	-0.14	-0.14	-0.16	-0.08	-0.04
	(0.06)**	(0.06)**	(0.06)**	(0.07)**	(0.03)**	(0.05)
Firm Age	-0.02	-0.02	-0.02	-0.01	-0.02	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
LOG(Value-added)	0.16	0.16	0.15	-0.01	-0.54	-0.35
	(0.41)	(0.41)	(0.42)	(0.46)	(0.43)	(0.46)
LOG(Sales)	0.97	0.97	0.99	1.13	1.74	1.43
	(0.44)**	(0.44)**	(0.44)**	(0.48)**	(0.45)***	(0.49)***
Input-Output Ratio	-0.54	-0.54	-0.54	-0.29	-0.20	-0.35
	(0.22)**	(0.22)**	(0.22)**	(0.31)	(0.20)	(0.24)
Export-Sales Ratio	-0.64	-0.63	-0.6	-0.59	-0.85	-0.72
	(0.61)	(0.61)	(0.61)	(0.66)	(0.60)	(0.64)
	County-Year Characteristics					
LOG(gdp/pop)	-0.4	-0.4	-0.4	-0.4	NO	NO
	(0.10)***	(0.10)***	(0.10)***	(0.10)***		
LOG(Pop)	-0.67	-0.67	-0.67	-0.67	NO	NO
	(0.15)***	(0.15)***	(0.15)***	(0.15)***		
Fiscal Pressure	0.07	0.07	0.06	0.06	NO	NO
	(0.05)	(0.05)	(0.05)	(0.05)		
Sample size	76,799	76,799	76,799	76,799	139,261	139,261
R^2	0.17	0.17	0.17	0.17	0.14	0.28

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the province-by-(1-digit)industry level (155 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.
2. In the prefecture-industry and county-industry fixed effects, industries are identified at the 1-digit level.

Table 3: Tenure Effect on the Change of VAT Enforcement

	Dependent Variable: Effective VAT Rate			
	Panel A: Baseline DD		Panel B: Tenure Effect	
	(1)	(2)	(3)	(4)
Post × Agr	2.47	2.47	0.64	0.84
	(0.81)***	(1.03)**	(1.21)	(1.27)
Post × Agr × Tenure	NO	NO	0.58	0.46
			(0.28)**	(0.26)*
Post × Agr × Tenure × CapInt	NO	NO	-1.31	-1.19
			(0.48)***	(0.45)***
Post × Agr × CapInt	NO	NO	1.20	0.78
			(1.90)	(1.87)
	Industrial Trends			
CapInt × Year	NO	NO	-0.37	-0.31
			(0.14)***	(0.14)**
	Fixed Effects			
County FE	YES	YES	NO	NO
County-(2)Indu FE	NO	NO	YES	YES
	Other Diff-in-Diff Terms			
Post	-0.05	-0.05	0.03	0.02
	(0.03)	(0.04)	(0.05)	(0.04)
Post × CapInt	NO	NO	-0.26	-0.43
			(0.79)	(0.74)
Agr × CapInt	NO	NO	-2.5	-2.12
			(0.90)***	(0.84)**
Agr × Tenure	NO	NO	-0.12	-0.12
			(0.20)	(0.21)
CapInt × Tenure	NO	NO	-0.07	-0.06
			(0.08)	(0.07)
Post × CapInt × Tenure	NO	NO	0.22	0.22
			(0.10)**	(0.09)**
Agr × CapInt × Tenure	NO	NO	0.35	0.41
			(0.28)	(0.28)
CapInt	NO	NO	738.82	613.87
			(279.83)***	(271.61)**
CapInt Square	NO	NO	-0.13	0.23
			(0.29)	(0.26)
Tenure Square	NO	NO	0.00	-0.01
			(0.01)	(0.01)

(Continued on next page)

Table 3: Tenure Effect on the Change of VAT Enforcement (Continued)

	2-Digit Industry Characteristics			
	(1)	(2)	(3)	(4)
Ownership	NO	NO	NO	-1.15
				(0.18)***
Profitability	NO	NO	NO	0.03
				(0.00)***
Firm Age	NO	NO	NO	0.04
				(0.01)***
LOG(V-added)	NO	NO	NO	-4.04
				(0.26)***
LOG(Sales)	NO	NO	NO	3.97
				(0.24)***
Input-Output Ratio	NO	NO	NO	0.03
				(0.01)**
Export-Sales Ratio	NO	NO	NO	-0.01
				-0.01
Sample size	17,964	17,964	158,885	158,762
R^2	0.41	0.41	0.54	0.57
Clustering	Prefecture (362 groups)	Province (31 groups)	Province-(1)Industry (135 groups)	Province-(1)Industry (135 groups)

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the province-by-(1-digit)industry level (155 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.
2. In the county-industry fixed effects, industries are identified at the 2-digit level.
3. In the clustering, province-industry refers to the 1-digit industry level.

Table 4: Tenure Effect on Profit Gap Rate

	Dependent Variable: Profit Gap Rate					
	Panel A: Baseline			Panel B: Control for *Tenure* × *Other Characteristics*		
	(1)	(2)	(3)	(4)	(5)	(6)
Tenure	-0.52	-0.99	-0.97	-13.86	-9.13	-9.21
	(0.27)*	(0.33)***	(0.33)***	(3.71)***	(2.27)***	(2.50)***
Tenure	NO	2.58	2.51	3.81	2.67	2.42
× CapInt		(0.79)***	(0.79)***	(0.87)***	(0.62)***	(0.58)***
	Industrial Trends					
CapInt * Year	NO	NO	1.19	1.16	1.18	1.27
			(0.65)*	(0.65)*	(0.61)*	(0.62)**
	Fixed Effects					
Province-Year FE	YES	YES	YES	YES	YES	YES
Prefecture-(1)Indu FE	YES	YES	YES	YES	YES	NO
County-(1)Indu FE	NO	NO	NO	NO	NO	YES
	Secretary's Controls					
Dummy (Duration < 5)	-0.43	-0.43	-0.41	-0.27	-0.37	-0.87
	(0.75)	(0.75)	(0.75)	(0.76)	(0.62)	(0.64)
Dummy (Duration > 5)	0.56	0.56	0.56	0.6	1.47	1.05
	(0.84)	(0.84)	(0.84)	(0.85)	(0.67)**	(0.73)
Tenure Square	0.1	0.1	0.1	0.11	0.03	0.02
	(0.05)**	(0.05)**	(0.05)**	(0.05)**	(0.04)	(0.04)
	Tenure × *Other Characteristics*					
	Not Included			*Included*		
	(1)	(2)	(3)	(4)	(5)	(6)
Tenure				2.72	2.34	2.5
× Ownership				(1.65)	(1.02)**	(1.11)**
Tenure				0.14	0.62	0.53
× Profitability				(0.37)	(0.33)*	(0.32)*
Tenure				0.03	0.02	0.03
× Firm Age				(0.03)	(0.02)	(0.02)
Tenure				0.43	-0.22	0.09
× LOG(Value-added)				(0.57)	(0.38)	(0.39)
Tenure				0.2	0.69	0.38
× LOG(Sales)				(0.51)	(0.33)**	(0.32)
Tenure				5.34	1.52	1.93
× Input-Output Ratio				(2.04)***	(0.98)	(0.91)**
Tenure				0.33	0.66	0.48
× Export Ratio				(0.74)	(0.52)	(0.49)

(Continued on next page)

Table 4: Tenure Effect on Profit Gap Rate (Continued)

	2-Digit Industry Characteristics					
	(1)	(2)	(3)	(4)	(5)	(6)
CapInt	-59.74	-64.31	-2438.69	-2390.39	-2435.3	-2596.31
	(8.30)***	(8.42)***	(1,308.48)*	(1,301.27)*	(1,215.51)**	(1,245.27)**
CapInt Square	20.04	20.03	20.26	20.47	19.12	19.28
	(8.29)**	(8.31)**	(8.32)**	(8.27)**	(6.97)***	(7.12)***
Ownership	1.72	1.78	1.79	-3.15	-0.33	-1.45
	(3.54)	(3.57)	(3.56)	(5.53)	(3.71)	(4.16)
Profitability	0.54	0.54	0.55	0.41	-0.55	-0.47
	(0.47)	(0.47)	(0.46)	(0.39)	(0.45)	(0.51)
Firm Age	0.09	0.09	0.09	0.03	0.04	-0.04
	(0.09)	(0.09)	(0.09)	(0.10)	(0.09)	(0.08)
LOG(Value-added)	-1.05	-1.04	-1.01	-1.63	1.22	-0.43
	(1.99)	(1.98)	(1.98)	(2.13)	(1.80)	(1.91)
LOG(Sales)	-2.87	-2.87	-2.9	-3.38	-5.28	-3.01
	(2.06)	(2.04)	(2.05)	(2.15)	(1.77)***	(1.82)*
Input-Output Ratio	6.51	6.48	6.48	1.11	4.88	4.08
	(2.03)***	(2.02)***	(2.03)***	(2.42)	(2.02)**	(1.68)**
Export-Sales Ratio	4.39	4.29	4.24	3.75	3.97	3.43
	(2.07)**	(2.07)**	(2.07)**	(2.11)*	(1.91)**	(2.01)*
	County-Year Characteristics					
LOG(gdp/pop)	0.34	0.35	0.35	0.34	NO	NO
	(0.46)	(0.46)	(0.46)	(0.46)		
LOG(Pop)	1.32	1.32	1.32	1.32	NO	NO
	(0.54)**	(0.54)**	(0.54)**	(0.54)**		
Fiscal Pressure	-0.03	-0.03	-0.02	-0.03	NO	NO
	(0.27)	(0.27)	(0.27)	(0.27)		
Sample size	76,799	76,799	76,799	76,799	139,261	139,261
R^2	0.2	0.2	0.2	0.2	0.17	0.3

Notes:
1. Ordinary least squares. Robust standard errors are clustered at the province-by-(1-digit)industry level (155 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.
2. In the prefecture-industry and county-industry fixed effects, industries are identified at the 1-digit level.

Table 5: Tenure Effect on Debt Ratio

	Dependent Variable: Debt Ratio					
	Panel A: Baseline			Panel B: Control for *Tenure × Other Characteristics*		
	(1)	(2)	(3)	(4)	(5)	(6)
Tenure	-0.33	-0.52	-0.5	-6.69	-3.69	-3.23
	(0.18)*	(0.20)**	(0.20)**	(1.80)***	(1.26)***	(1.26)**
Tenure		1.07	0.94	1.22	0.81	0.51
× CapInt		(0.33)***	(0.28)***	(0.41)***	(0.31)**	(0.33)
	Industrial Trends					
CapInt × Year			2.09	2.1	1.89	1.79
			(0.50)***	(0.50)***	(0.51)***	(0.52)***
	Fixed Effects					
Province-Year FE	YES	YES	YES	YES	YES	YES
Prefecture-(1)Indu FE	YES	YES	YES	YES	YES	NO
County-(1)Indu FE	NO	NO	NO	NO	NO	YES
	Secretary's Controls					
Dummy (Duration < 5)	1.04	1.05	1.07	1.12	0.46	0.30
	(0.41)**	(0.41)**	(0.41)**	(0.41)***	(0.35)	(0.36)
Dummy (Duration > 5)	1.5	1.5	1.51	1.52	0.41	0.62
	(0.56)***	(0.56)***	(0.56)***	(0.56)***	(0.44)	(0.49)
Tenure Square	0.07	0.07	0.07	0.06	-0.03	-0.04
	(0.03)**	(0.03)**	(0.03)**	(0.03)*	(0.02)	(0.03)
	Tenure × Other Characteristics					
	Not Included			*Included*		
	(1)	(2)	(3)	(4)	(5)	(6)
Tenure				2.03	1.37	1.2
× Ownership				(0.77)***	(0.55)**	(0.58)**
Tenure				0.04	0.17	0.36
× Profitability				(0.21)	(0.13)	(0.14)**
Tenure				0.02	0.00	0.01
× Firm Age				(0.02)	(0.02)	(0.02)
Tenure				1.04	0.61	0.78
× LOG(Value-added)				(0.30)***	(0.21)***	(0.22)***
Tenure				-0.61	-0.3	-0.48
× LOG(Sales)				(0.29)**	(0.21)	(0.22)**
Tenure				1.26	-0.14	-0.24
× Input-Output Ratio				(0.91)	(0.56)	(0.51)
Tenure				-0.21	0.23	0.27
× Export Ratio				(0.45)	(0.38)	(0.38)

(Continued on next page)

Table 5: Tenure Effect on Debt Ratio (Continued)

	2-Digit Industry Characteristics					
	(1)	(2)	(3)	(4)	(5)	(6)
CapInt	10.2	8.31	-4180.6	-4191	-3786.52	-3583.14
	(3.22)***	(3.31)**	(1,006.74)***	(1,006.41)***	(1,026.71)***	(1,033.22)***
CapInt Square	-15.84	-15.84	-15.43	-15.43	-9.71	-9.02
	(3.39)***	(3.37)***	(3.39)***	(3.40)***	(2.64)***	(2.73)***
Ownership	-18.37	-18.34	-18.33	-22.02	-19.68	-19.04
	(2.52)***	(2.52)***	(2.50)***	(2.86)***	(2.27)***	(2.51)***
Profitability	-0.13	-0.14	-0.12	-0.15	-0.14	-0.28
	(0.22)	(0.23)	(0.23)	(0.24)	(0.08)	(0.13)**
Firm Age	-0.02	-0.02	-0.01	-0.05	-0.01	-0.04
	(0.06)	(0.06)	(0.06)	(0.08)	(0.07)	(0.07)
LOG(Value-added)	-3.32	-3.32	-3.27	-5.06	-4.53	-5.42
	(1.07)***	(1.06)***	(1.07)***	(1.28)***	(1.00)***	(1.01)***
LOG(Sales)	3.46	3.46	3.4	4.43	4.49	5.26
	(1.07)***	(1.07)***	(1.07)***	(1.29)***	(0.99)***	(1.01)***
Input-Output Ratio	2.61	2.6	2.61	1.36	4.05	3.13
	(1.67)	(1.67)	(1.66)	(1.71)	(1.84)**	(1.54)**
Export-Sales Ratio	0.71	0.67	0.59	0.94	0.49	0.22
	(1.09)	(1.1)	(1.11)	(1.270	(1.17)	(1.15)
	County-Year Characteristics					
LOG(gdp/pop)	-0.76	-0.75	-0.76	-0.75	NO	NO
	(0.32)**	(0.32)**	(0.32)**	(0.32)**		
LOG(Pop)	-1.56	-1.56	-1.57	-1.57	NO	NO
	(0.35)***	(0.35)***	(0.35)***	(0.35)***		
Fiscal Pressure	0.08	0.08	0.09	0.09	NO	NO
	(0.11)	(0.11)	(0.11)	(0.11)		
Sample size	76,798	76,798	76,798	76,798	139,259	139,259
R^2	0.12	0.12	0.12	0.12	0.12	0.28

Notes:

1. Ordinary least squares. Robust standard errors are clustered at the province-by-(1-digit)industry level (155 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

2. In the prefecture-industry and county-industry fixed effects, industries are identified at the 1-digit level.

Table 6: Tenure Effect on ETC

	Dependent Variable: ETC Ratio				
	(1)	(2)	(3)	(4)	(5)
Tenure	-0.03	-0.22	-0.24	-0.22	-0.22
	(0.02)	(0.28)	(0.28)	(0.27)	(0.28)
Tenure	0.09	0.09	0.09	0.09	0.09
× CapInt	(0.01)***	(0.09)***	(0.01)***	(0.01)***	(0.01)***
	Fixed Effects				
Prov-(1)Indu FE	YES	YES	YES	YES	YES
	Secretary's Controls				
Dummy (Duration < 5)	NO	NO	0.01	NO	0.01
			(0.11)		(0.10)
Dummy (Duration > 5)	NO	NO	-0.04	NO	-0.03
			(0.16)		(0.16)
Dummy (Age < 50)	NO	NO	NO	-0.07	-0.07
				(0.17)	(0.16)
Dummy (Age >= 50)	NO	NO	NO	0.7	0.72
				(0.67)	(0.65)
	Tenure × Other Characteristics				
	(1)	(2)	(3)	(4)	(5)
Tenure	NO	0.07	0.08	0.08	0.08
× Ownership		(0.15)	(0.15)	(0.15)	(0.15)
Tenure	NO	-0.13	-0.13	-0.13	-0.13
× Profitability		(0.13)	(0.13)	(0.13)	(0.13)
Tenure	NO	0.00	0.00	0.00	0.00
× Firm Age		(0.00)	(0.00)	(0.00)	(0.00)
Tenure	NO	0.00	0.00	0.00	0.00
× Manager Tenure		(0.01)	(0.01)	(0.01)	(0.01)
Tenure	NO	-0.01	0.00	-0.01	-0.01
× LOG(Sales)		(0.05)	(0.05)	(0.05)	(0.05)
Tenure	NO	0.02	0.02	0.03	0.03
× LOG(V-added)		(0.05)	(0.05)	(0.05)	(0.05)

(Continued on next page)

Table 6: Tenure Effect on ETC (Continued)

	2-Digit Industry Characteristics				
	(1)	(2)	(3)	(4)	(5)
CapInt	0.47	0.51	0.5	0.17	0.16
	(0.42)	(0.42)	(0.42)	(0.62)	(0.61)
Ownership	0.06	-0.05	-0.07	-0.08	-0.09
	(0.24)	(0.34)	(0.34)	(0.34)	(0.34)
Profitability	-0.01	0.00	0.00	0.00	0.00
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Firm Age	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Manger Tenure	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
LOG(Sales)	-0.4	-0.38	-0.38	-0.37	-0.37
	(0.07)***	(0.14)**	(0.14)**	(0.14)**	(0.14)**
LOG(V-added)	0.19	0.15	0.15	0.14	0.14
	(0.07)**	(0.13)	(0.13)	(0.14)	(0.14)
Sample size	1,270	1,270	1,270	1,270	1,270
R^2	0.33	0.33	0.33	0.34	0.34

Notes:

1. Ordinary least squares. Robust standard errors are clustered at the province-by-(1-digit)industry level (155 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

2. In the province-industry fixed effects, industries are identified at the 1-digit level.

Table 7: Return to ETC: Effective VAT Rate

	Dependent Variable: Effective VAT Rate							
	Panel A: Baseline	Panel B: Control for *(ETC / 100) × Other Characteristics*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ETC / 100	0.74	1.36	0.95	0.56	0.93	1.54	1.56	1.41
	(0.19)***	(0.71)*	(0.28)***	(0.24)**	(0.35)***	(0.75)**	(0.90)*	(1.22)
ETC / 100	-0.48	-0.48	-1.44	-0.46	-0.48	-0.58	-0.59	-1.53
× CapInt	(0.18)**	(0.19)**	(0.57)**	(0.19)**	(0.18)***	(0.22)**	(0.24)**	(0.57)***
	Fixed Effects							
Prefecture-(1)Indu FE	YES	YES	YES	YES	YES	YES	YES	YES
	(ETC / 100) × Other Characteristics							
ETC /100		-0.63						-0.28
× Ownership		(0.67)						(0.77)
ETC /100			-0.62					-0.67
× Profitability			(0.36)*					(0.37)*
ETC /100				0.02				0.01
× Firm Age				(0.01)				(0.01)
ETC /100					-0.03			-0.01
× Manager Tenure					(0.03)			(0.04)
ETC /100						-0.08		-0.47
× LOG(Sales)						(0.07)		(0.56)
ETC /100							-0.09	0.49
× LOG(V-added)							(0.09)	(0.57)
	2-Digit Industry Characteristics							
CapInt	0.09	0.09	0.11	0.09	0.09	0.09	0.09	0.11
	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***
Ownership	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Profitability	0.01	0.01	0.03	0.01	0.01	0.01	0.01	0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Firm Age	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Manager Tenure	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
LOG(Sales)	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18
	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***
LOG(V-added)	-0.18	-0.18	-0.18	-0.18	-0.18	-0.18	-0.18	-0.19
	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***	(0.03)***
Constant	-0.08	-0.09	-0.09	-0.08	-0.09	-0.09	-0.09	-0.09
	(0.05)*	(0.05)*	(0.05)*	(0.05)*	(0.05)*	(0.05)*	(0.05)*	(0.05)*
Sample size	11,291	11,291	11,291	11,291	11,291	11,291	11,291	11,291
R^2	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13

Notes:

1. Ordinary least squares. Robust standard errors are clustered at the province-by-(1-digit)industry level (155 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.
2. In the prefecture-industry fixed effects, industries are identified at the 1-digit level.
3. In this table, ETC = entertainment and travel expenditure / total sales. That is, ETC here is scaled down by dividing 100 on its standard measurement used in other regressions.

Table 8: Return to ETC: Collateral Requirement for Bank Loan

	Dependent Variable: Collateral Dummy							
	Panel A: Baseline	Panel B: Control for *(ETC / 100) × Other Characteristics*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ETC / 100	0.64	0.49	0.62	0.73	0.88	1.43	0.92	1.68
	(0.28)**	(0.77)	(0.28)**	(0.38)*	(0.36)**	(0.99)	(0.83)	(1.94)
ETC / 100	-0.62	-0.62	-0.61	-0.63	-0.63	-0.72	-0.66	-0.71
× CapInt	(0.27)**	(0.27)**	(0.27)**	(0.27)**	(0.27)**	(0.31)**	(0.31)**	(0.32)**
	Fixed Effects							
Prefecture-(1)Indu FE	YES	YES	YES	YES	YES	YES	YES	YES
	(ETC / 100) × Other Characteristics							
ETC / 100		0.15						0.03
× Ownership		(0.78)						(1.18)
ETC /100			0.04					0.07
* Profitability			(0.07)					(1.19)
ETC / 100				-0.01				-0.01
× Firm Age				(0.02)				(0.03)
ETC / 100					-0.04			-0.03
× Manager Tenure					(0.06)			(0.07)
ETC / 100						-0.08		-0.45
× LOG(Sales)						(0.09)		(0.33)
ETC / 100							-0.03	0.40
× LOG(V-added)							(0.09)	(0.28)
	2-Digit Industry Characteristics							
CapInt	0.12	0.12	0.12	0.12	0.12	0.13	0.12	0.13
	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.02)***
Ownership	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
	(0.03)**	(0.03)**	(0.03)**	(0.03)**	(0.03)**	(0.03)**	(0.03)**	(0.03)**
Profitability	0.01	0.01	0	0.01	0.01	0.01	0.01	0
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Firm Age	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Manager Tenure	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
LOG(Sales)	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.06
	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***
LOG(V-added)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	-0.26	-0.26	-0.26	-0.26	-0.26	-0.27	-0.26	-0.27
	(0.05)***	(0.05)***	(0.05)***	(0.05)***	(0.05)***	(0.05)***	(0.05)***	(0.06)***
Sample size	11,291	11,291	11,291	11,291	11,291	11,291	11,291	11,291
R^2	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16

Notes:

1. Ordinary least squares. Robust standard errors are clustered at the province-by-(1-digit)industry level (155 groups). *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

2. In the prefecture-industry fixed effects, industries are identified at the 1-digit level.

3. In this table, ETC = entertainment and travel expenditure / total sales. That is, ETC here is scaled down by dividing 100 on its standard measurement used in other regressions.

Chapter 3

VAT Rate Dispersion and Aggregate Efficiency in China*

Abstract

This paper studies dispersion in the effective VAT rate across manufacturing firms in China and assesses its impact on aggregate production efficiency from 2000 to 2007. Using a structural model based on Hsieh and Klenow (2009), I find that a tax-neutral reform which eliminates the dispersion in VAT rates produces a gain in aggregate TFP in the order of 7.9% of GDP on average in the period from 2000 to 2007.

*The author is grateful to Francesco Caselli for his supervision on this project. Acknowledgements go to Alwyn Young, Tim Besley, Johannes Spinnewijn, Ximin Yue, Liutang Gong, John Whalley for their suggestions and comments. The author also appreciates the feedback from participants in Work-in-progress Seminars on Macroeconomics, Public Economics, Development and Growth at LSE, and in Seminars at Renmin University of China, Beijing Normal University. This project is funded by the International development Research Council in Canada. The author bears responsibility for any errors within the paper.

1 Introduction

Non-lump-sum taxation leads to distortions and efficiency loss. Typically, the welfare loss can be captured by the Harberger triangle created by the tax wedge between the supply curve and the demand curve (Harberger, 1964a, b). In this paper, I look into another situation where tax rate dispersions across monopolistic competitive firms may also generate an efficiency loss via the channel of resource misallocation.

The mechanism of misallocation has been put under the spotlight in quite a few recent influential macro-development studies (Hsieh and Klenow, 2009; Rogerson and Rustuccia, 2008), yet the sources of distortion and misallocation is not clearly identified and directly measured. This paper, in an attempt to fill this gap in knowledge, proposes the effective VAT rate as a direct measurement of one kind of distortion and show its dispersions across firms could lead to significant loss in aggregate production efficiency. As a result, the findings could hopefully shed light on the VAT reform agenda in the search for a efficiency-enhancing VAT system as advocated by Mirrlees *et al.* (2011).

The mechanism underlying the impact of tax rate dispersion on efficiency loss is intuitive: the firm paying a low tax rate is able to steal business from its high-tax rivals and therefore results in resource misallocation between firms – low-tax firms are bigger than they ought to be and high-tax firms are smaller compared with the optimal firm size distribution which maximizes the aggregate productivity.

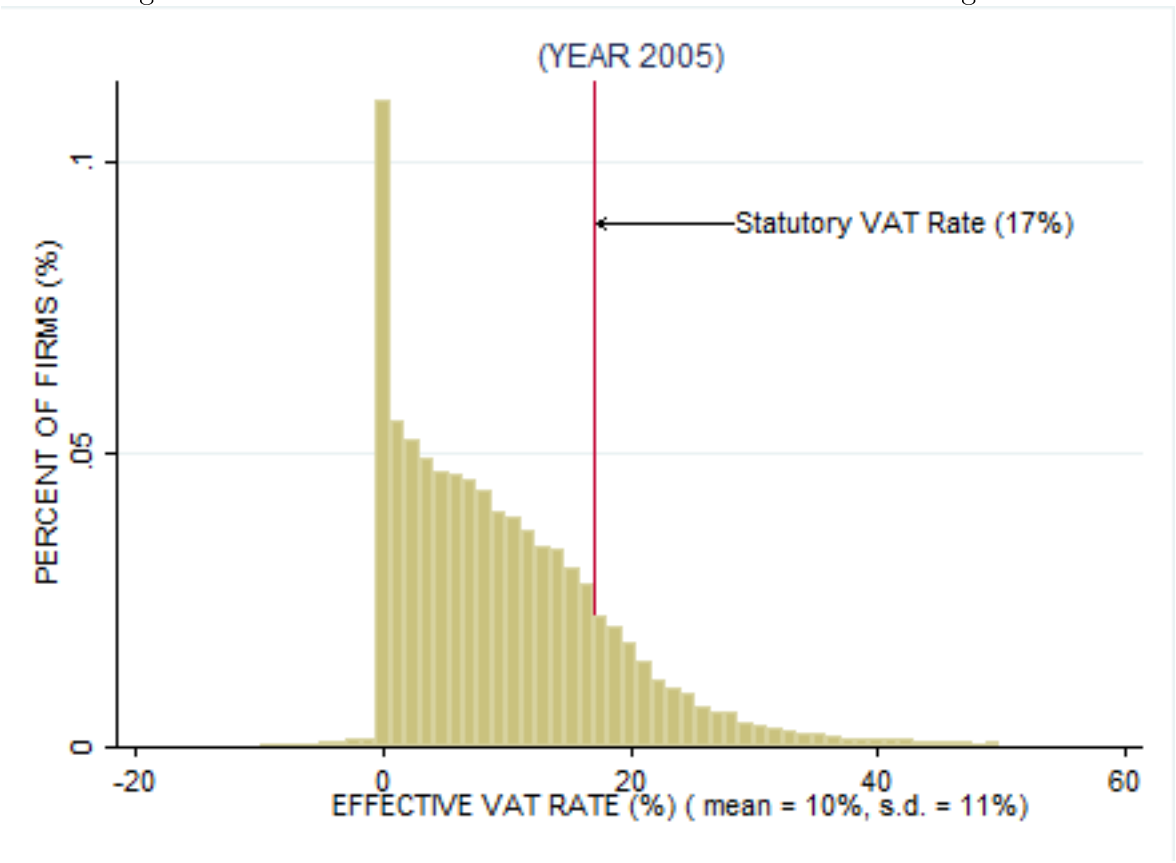
As will be demonstrated in the later part of this paper, the VAT rate dispersion in China's manufacturing firms is not only large in magnitude, but is also persistent, coming from within four-digit industry and almost uncorrelated with some important firm's characteristics. These attributes allow us to gauge its impact on efficiency loss in a theoretical structural model based on Hsieh and Klenow (2009). I find that a tax-neutral reform which eliminates the dispersion in VAT rates produces a gain in aggregate TFP in the order of 7.9% of GDP on average in the period from 2000 to 2007.

Figure 1.1 gives us a first glimpses at the distribution of the effective VAT rate across China's manufacturing firms, where the effective VAT rate is defined as the ratio of payable VAT to the value-added of a firm. Although the standard statutory tax rate is 17%, the effective tax rate varies a lot, with a standard deviation of 11% around the mean 10%. The majority of firms pay an actual effective tax rate that is lower than the legitimate level.[1]

In this paper, I only focus on VAT in China, although one may also consider the efficiency loss caused by various types of taxation, such as sales tax, capital gains tax,

[1]It is a bit striking that the effective VAT rate in Figure 1.1 does not stay around the standard statutory tax rate 17%. But apart from the statutory reasons such as differential tax rates and export tax refund, that is actually a vivid indication of poor VAT administration in China. Empirical studies suggest that the VAT administration in China can be affected by shocks on local governments budget (Chen, 2015a) and vary over local political cycles (Chen, 2015b).

Figure 1.1: VAT Rate Distribution over China's Manufacturing Firms



personal income tax, resource tax, and others. There are three reasons for my preference to VAT. First, it is VAT, rather than the corporate income tax, that is particularly relevant to the distortion on product in the theoretical model. Second, VAT is an important tax in China. It accounts for nearly one-third of the government total tax revenue, and therefore has gigantic welfare implication. Third, effective VAT rate at the firm level is available in China's firm survey.

This paper is organized as follows. Section 2 introduces a toy model that explains the underlying mechanism of misallocation caused by dispersion in effective VAT rate across firms. Section 3 reviews the related literature and points out contributions of this paper. Section 4 briefly introduces the institutional background of VAT in China. Section 5 presents a theoretical framework that relates tax rate dispersion to aggregate production efficiency loss. Section 6 summarizes the data used in this paper. Section 7 reports the results and main findings. Several robustness checks are conducted in Section 8, and Section 9 is the conclusion.

2 Conceptual Framework

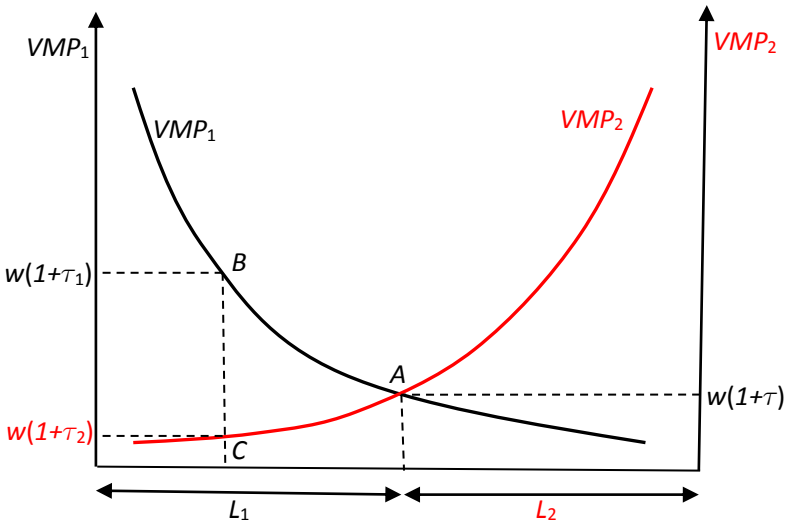
To understand the working mechanism of Hsieh and Klenow (2009), let us first consider a model economy consisting of two monopolistic competitive firms with one production factor (L units in total). The output Y is assumed to be produced with constant-return-to-scale technology: $Y_i = A_i L_i$, $i = 1, 2$, where A is the physical productivity (TFP). The monopolistic competition between two firms is featured with the output aggregator in Dixit-Stiglitz fashion: $Y = \left(Y_1^{\frac{\sigma-1}{\sigma}} + Y_2^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$, where σ is the elasticity of substitution, and it also reflects the degree of competition between firms. The GDP of this economy is aggregated over the value-added of all firms: $PY = P_1 Y_1 + P_2 Y_2$, where the final output Y can be chosen as the numeraire and P is normalized to 1.

Monopolistic competition implies that each firm faces a downward sloping demand curve for its product: $P_i = P\left(\frac{Y}{Y_i}\right)^{1/\sigma}$. The revenue (or the value-added) of each firm can be written as $P_i Y_i = PY^{\frac{1}{\sigma}} Y_i^{\frac{\sigma-1}{\sigma}} = PY^{\frac{1}{\sigma}} (A_i L_i)^{\frac{\sigma-1}{\sigma}}$, which is a function of input L_i. Suppose the profit maximizing firm faces the firm-specific tax rate on factor income wL_i. Maximizing profit $P_i Y_i - w(1+\tau_i)L_i$, $i = 1, 2$ over L_i gives rise to the first-order condition: $VMP_{L_i}(L_i) = w(1+\tau_i)$, where the value of marginal product $VMP_{L_i} = \frac{\partial(P_i Y_i)}{\partial L_i}$. Since VMP_{L_i} is positively proportional to $A_i^{\frac{\sigma-1}{\sigma}} L_i^{-\frac{1}{\sigma}}$, it is downward sloping in L_i. Suppose a benevolent government attempts to achieve the aggregate production efficiency by maximizing GDP subject to the monopolistic competitive market structure and resource constraint $L_1 + L_2 = L$. Obviously, the optimality condition requires $VMP_{L_1} = VMP_{L_2}$ (or equivalently, equal revenue productivity $TFPR_i$, as $TFPR_i = \frac{P_i Y_i}{L_i} = \frac{\sigma}{\sigma-1} VMP_{L_i}$). In a decentralized economy, the production efficiency is achieved if, and only if, there are no tax rate dispersions across firms; that is, $\tau_1 = \tau_2 = \tau$. It could also be put another way: the tax rate dispersion $var(\tau_i)$ leads to *additional* productivity loss aside from the welfare loss caused by the distortionary tax τ itself.

In this model economy, the efficiency loss is caused by the resource misallocation between monopolistic competitive firms. To be specific, the loss comes from the fact that the low-tax firm is able to charge a lower price on its product, thus stealing business from its high-tax rival. The magnitude of efficiency loss therefore relies on two factors: (1) the dispersion of tax rate across firms and (2) the elasticity of substitution between the firms' products.

The idea above can be illustrated in Figure 2.1. The social optimal allocation is determined by the intersection point between VMP_1 and VMP_2. In a decentralized economy, this can be achieved by setting $\tau_1 = \tau_2$. If $\tau_1 \neq \tau_2$, monopolistic competition gives rise to allocation where $VMP_1 \neq VMP_2$. In the case of $\tau_1 > \tau_2$, high-tax firm (or low-tax firm) is smaller (or bigger) than the social optimal size, as is referred to as

Figure 2.1: Tax Rate Dispersion and Production Efficiency Loss



resource misallocation. Although the resource allocation is Pareto optimal[2], potential gain in GDP/Productivity could still be achieved by re-allocating a certain amount of L from firm 2 to firm 1. The production efficiency loss can be measured by the area of $\Delta ABC \simeq \frac{1}{2}BC \times (BC \times ctg(\angle A)) = \frac{1}{2}BC^2 ctg(\angle A)^3$, where $BC^2 = [w(\tau_1 - \tau_2)]^2 \propto var(\tau)$, and $ctg(\angle A)$ ~slope of $VMP_{L_i} \sim \sigma$, where "~" denotes "to be positively proportional to".

In a more general setting with more than two firms, we can go even further with particular assumptions and prove $logTFP = logTFP^e - \frac{\sigma-1}{2} var\left(\log(1+\tau_i)\right)$, where TFP is the aggregate total factor productivity, and TFP^e is the TFP without tax rate dispersion. Intuition is: (1) Bigger $var(\tau)$ $\Rightarrow$ more tax rate dispersion $\Rightarrow$ more resource mis-allocation $\Rightarrow$ more efficiency loss; (2) Higher $\sigma \Rightarrow$ products are more substitutable $\Rightarrow$ more business-stealing by low-tax firm $\Rightarrow$ more resources misallocated from high-tax firm to low-tax firm $\Rightarrow$ more efficiency loss.

At this juncture, a query arise regarding the relationship between the problem of GDP maximization set out above, and the theory of optimal taxation. To see this, I suppose $\left(Y_1^{\frac{\sigma-1}{\sigma}} + Y_2^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$ is actually the social welfare function (or utility function of homogeneous consumers). In a standard Ramsey optimal taxation setting, the problem of government can be written as

$$\max_{\{L_1, L_2, \tau_1, \tau_2\}} \left((A_1 L_1)^{\frac{\sigma-1}{\sigma}} + (A_2 L_2)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{2.1}$$

[2]Since we can not increase Y_1 without reducing Y_2, given the resource constraint on L.

[3]In our analysis, I ignore the simultaneous upward shift of VMP_i $(i = 1, 2)$ curves because Y increases when the disparity between τ_1 and τ_2 are eliminated. This simplification makes the implication of ΔABC more straightforward.

$$s.t. \begin{cases} L_1 \text{ and } L_2 \text{ in monopolistic competitive equilibrium} \\ (\tau_1 L_1 + \tau_2 L_2) w \geqslant T \end{cases}$$

where T is the tax revenue requirement of the government. In partial equilibrium where w is taken as exogeneous to Ramsey problem (1), one can prove the optimal taxation is characterized by $\tau_1 = \tau_2 = \tau$.

3 Related Literature

This paper is mainly related both to misallocation and to efficiency loss of taxation. It also sheds some lights on issues of VAT administration.

3.1 Misallocation and Production Efficiency

This paper follows a strand of voluminous "development accounting" literature over last two decades. These items of literature have been comprehensively reviewed by Caselli (2005) and Hsieh and Klenow (2010), so I do not attempt to provide a redundant list of related papers. However, in order to set out precisely where I start, I would briefly refer to several papers in particular.

Over the past five years, there has been a sharp increase in research interests focusing on the firm-level distortions and their impact on TFP. In Hsieh and Klenow (2009), firm specific distortions result in dispersion in revenue productivity across firms and lead to loss in aggregate TFP. Restuccia and Rogerson (2008) introduce the distortions in the Lucas span-of-control model and find enormous aggregate TFP loss if distortions and firm productivity are positively correlated. Guner, Ventura, Xu (2008) show that the policies that distort the size of production units could lead to reductions in output and output per establishment up to 8.1% and 25.6% respectively. Caselli and Gennaioli (2005) propose a model in which the poor contract enforcement in dynastic family firms may be a substantial contributor to observed cross-country differences in aggregate TFP between rich and poor countries. Bartelsman, Haltiwanger, and Scarpetta (2009) investigate the effect of firm-specific distortions on aggregate performance in the cross-country context. Jones (2011) shows how the sectoral distortions can be amplified by the input-output production structure.

In particular, our study moves in the same direction as Hsieh and Klenow (2009), providing a structural model in an attempt to measure the efficiency loss caused by various distortions. In Hsieh and Klenow (2009) however, the distortions are subject to two deficiencies: (1) The distortions are not directly observable; we are not aware of the sources of distortions. In their study, the model embraces all distortions in the capital, including

labour, output markets, taxes, regulations, market frictions, and many other unobservable factors; a long list which is difficult to examine one by one in quantitative way. (2) The distortions, being extracted from a structural model, are obviously vulnerable to model setting and the parametric values of the model.

This paper only focuses on the VAT rate dispersion as a source distortion, which is able to overcome the shortcomings of Hsieh and Klenow (2009) mentioned above for two reasons. (1) effective VAT rate is directly observable at the firm level. (2) VAT rate is independent of the model setting.

In this paper, I find the VAT rate dispersion mainly comes from firms' time-persistent factors such as geographic location, ownership and internal organization. This finding may provide a partial answer to the question, raised by Moll and Banerjee (2010) and popular in cross-country development studies, of why the distortion persists.

3.2 Taxation and Efficiency loss

The impact of non-lump-sum tax on welfare is usually measured by the Harberger Triangle created by the tax wedge between the supply curve and the demand curve. The basic rationale in public finance theory argues that the tax produces a tax wedge between buyers and sellers, changes the relative prices of goods, results in the behavioural response of economic agents, destroys the conditions for Pareto optimality, and thus leads to a deadweight loss on welfare. According to this argument, the deadweight loss is proportional to the elasticity of behavioural response to the change in tax rate (the base of the Harberger Triangle (Hines, 1999)). A series of empirical studies, inspired by the seminal work of Arnold Harberger in the 1960s (Harberger, 1964a, b), have been made in order to estimate the welfare impact of a wide array of tax-induced distortions; including those to labour supply (Browing, 1975), saving (Feldstein, 1978), corporate taxation (Shoven, 1976), and the consumption of goods (King, 1983). The magnitudes of the welfare loss, nevertheless, are still ambiguous in these studies, ranging from as low as 2.5% of the tax revenue (Harberger, 1964b) to the stunning 200% (Feldstein, 1999).

This paper suggests an additional mechanism of resource misallocation through which tax rate dispersions may also result in efficiency loss. In a traditional Harberger triangle, economists actually studied the impact of the mean, or the first-order moment, of tax rates on economic efficiency. Tax rate dispersion, measured by its variance or the second-order moment, is set on the central stage of this paper, which could be viewed as a complement to the Harberger triangle. From a more comprehensive perspective, both the first-order moment and the second-order moment of the tax rate work together in the provision of a complete view on the efficiency loss of taxation.

This study also relates directly to the Production Efficiency Theorem proved in Diamond and Mirrlees (1971). In this paper we will observe that production efficiency can

be achieved when there is no tax rate dispersion.

3.3 VAT and Welfare Loss

VAT in practice usually leads to welfare loss, due to various reasons including exemption of some sectors out of the VAT system, differential tax rates and zero rating, and other similar reasons. Mirrlees *et al.* (2011, Chapter 9) finds that the government tax revenue could increase by three billion pounds by eliminating the differential tax rates, while still maintaining the welfare level of the UK family. Zee (2006) finds that the exemption of financial services out of the VAT system leads to a efficiency loss amounting to about 30% of the GDP of the UK. Piggott and Whalley (2001) analyse the efficiency loss caused by the enlarging of the informal sector after the VAT expansion in Canada.

This paper, in attempt to provide a new perspective on the efficiency loss caused by the imperfect VAT system in practice, is a complement to the existing studies on efficiency loss of VAT.

4 Institutional Background of VAT in China

China has introduced VAT as early as 1979. There was a fundamental reform in 1994 on the whole Chinese tax system, which transformed the Chinese VAT into modernity and made it consistent with those in developed countries. Since 1994, the VAT has been administered by the State Administration of Taxation (SAT). Despite several adjustments has been made over the last decade, its framework remained stable.

4.1 Taxpayers

There are two types of VAT taxpayers in China, acknowledged in terms of the criteria of turnover of sale goods and services, and the condition of their accounting system. One is *general taxpayer*, the other is *small taxpayer*[4]. In the firm survey dataset I use in this paper, all the firms are General VAT taxpayers.

4.2 Taxable items and tax rates

Tax rates of VAT in China are flat, with differential rates (17%, 13% and 0%) applied to different taxable goods and services. There are three main rates: (1) the standard rate

[4]The standard definition of Small Taxpayer is: (1) where the taxpayer is engaged in the production of goods or in the provision of taxable services as his sole or principal business and his annual turnover is less than one million RMB ¥; or (2) where the taxpayer is engaged in the wholesale or retail of goods and his annual turnover is less than 1.8 million RMB ¥. However, A Small Taxpayer who maintains a sound accounting system and is able to provide accurate accounting records for taxation purposes may be registered as a General Taxpayer.

at 17% for most taxable goods and services; (2) the reduced rate at 13% for agricultural goods, public utilities, agriculture production inputs (such as fertilizers and agricultural machinery), books, newspapers, magazines, and other products; (3) the zero rating for exports.

4.3 Tax Exemption and Refund

Items exempted from tax include: (1) some agricultural production materials; (2) contraceptive medicines and devices; (3) antique books purchased from the public; (4) some goods imported for direct use in scientific research, experiment and education, and (5) materials and equipment imported as donations, for poverty relief or other charity purposes.

In the case of the zero tax rate applicable to exported goods, the exporters may apply to the tax authorities for the input tax refund on those particular exported goods. At present, the refund rates vary, at 3%, 5%, 8%, 13% and 17%.

4.4 Tax Authorities

To accommodate the tax revenue sharing system after the 1994 reform, the tax organisations at and below the provincial level are divided into offices of the SAT (State Administration of Taxation) and local tax bureaus.

According to the rules of the State Council on revenue sharing system, the tax revenue in China may be divided into Central tax revenue, local tax revenue and the tax revenue shared between the Central and local governments. Here "local" refers to governments at and below the provincial level, including provincial, municipal, county and village. The tax revenue sharing system between central and local governments was clearly defined in the 1994 reform. Tax revenue shared between the Central and local governments is arranged in the following way: (1) For domestic VAT, 75% is retained by Central Government and 25% is retained by Local Governments. (2) VAT collected by Customs belongs to Central tax revenue.

4.5 Tax Compliance and Tax Administration

The degree of tax compliance varies across firms on the basis of firm ownership, firm size, access to financial market, firm location, and various other factors.

Cai and Liu (2009) show that, after controlling for other characteristics, firms facing higher tax rates or tighter financial constraints and smaller firms are more likely to avoid tax.

Another salient feature in China is that State-Owned Enterprises (SOE) have more incentive to comply with the tax code than their non-SOE counterparts. This is largely

due to the fact the head of the SOE is generally motivated by pursuing his own promotion as opposed to maximising the firm's post-tax profit. In China, the performance of the SOE is usually measured in terms of the total taxes and profits contributed to the government, while for the Non-SOE, a penny paid as tax is considered to be a penny lost. Moreover, Non-SOE firms generally have many disadvantages including higher financing costs and a smaller scale of operation in any competition with SOE rivals. All these features of Non-SOE firms mean that they are less likely to adhere to the tax code.

In regard to tax administration, similar to many other developing countries, presumptive taxation and minimum tax are prevalent in China's taxation practice, especially for the lower tier tax authorities when levying taxes on small enterprises.[5] These practices are usually implemented in order to counter deficiencies in tax administration. Although the tax codes are not strictly enforced, these practices are justifiable in an economic sense, given the limited resources available (such as the number of tax staff, especially staff of superior quality; computers and other facilities) to the tax authorities, as well as the large number of small businesses unable to keep standard accounting books and to file a qualified tax declaration.

Presumptive methods may help reduce audit time and cost. A simple approach is to levy a lump sum on all businesses. A more sophisticated approach involves a census of taxpayers and an estimation of income, assets, turnover, or other alternatives. At the same time, the taxpayers can negotiate their tax liability with tax officials. Since presumptive tax generally comprises a tax on average or "normal" tax base, the marginal tax rate on a tax base above this average tax base is zero, leading to a declining effective average tax rate as the size of the actual tax base increases.

Minimum tax is another method for an imperfect tax administration to levy taxes by requiring minimum tax contributions. Generally, tax administration in China is inclined to focus on large taxpayers, while rudimentary presumptive methods or minimum tax have to be used for contributions from smaller taxpayers.

At the local (provincial, municipal, county) government level, the tax authorities (SAT and local tax bureau) regularly list the Major Tax Contributors among all the taxpayers each year according to a certain criterion (generally, their tax contribution from the previous year). Although the number of Major Tax Contributors in simple numerical terms account for a tiny part of all taxpayers, they contribute a much larger proportion of the tax revenue. For example, in the year 2007, about 70% of the tax revenue in the city of GuangZhou came from the Major Tax Contributors, who number less than 1% of the registered taxpayers in the city. Between 2001 and 2005, over 30% of the tax revenue in the city of ShenZhen was collected from the top 100 Major Tax Contributors, who

[5]For presumptive tax and minimum tax, refer respectively to Bulutoglu (1995) and Stotsky (1995) for details.

Table 1: VAT Rate Dispersions within (Four-digit) Industries

number less than 0.05% of the taxpayers.[6]

In addtion to the tax administrative practice mentioned above, evidences show that the VAT administration in China can be affected by incentives faced by the local governments and local politicians. The effective VAT rate may vary over the tightness of local governments budget (Chen, 2015a) and local political cycles (Chen, 2015b).

Because of the rugged tax codes (differential rates, zero rating, tax refund, etc.) as well as the poor tax administration mentoned above, the effective VAT rate across firms in China is quite dispersed. More importantly, Table 1 shows that most of the dispersion comes from within four-digit industries. This allows us to study the effects of tax rate dispersion on resource allocation among monopolistic firms in the same industry as the theory suggests.

Figure 4.1: VAT Rate Distribution over Firm in a Four-digit Industry

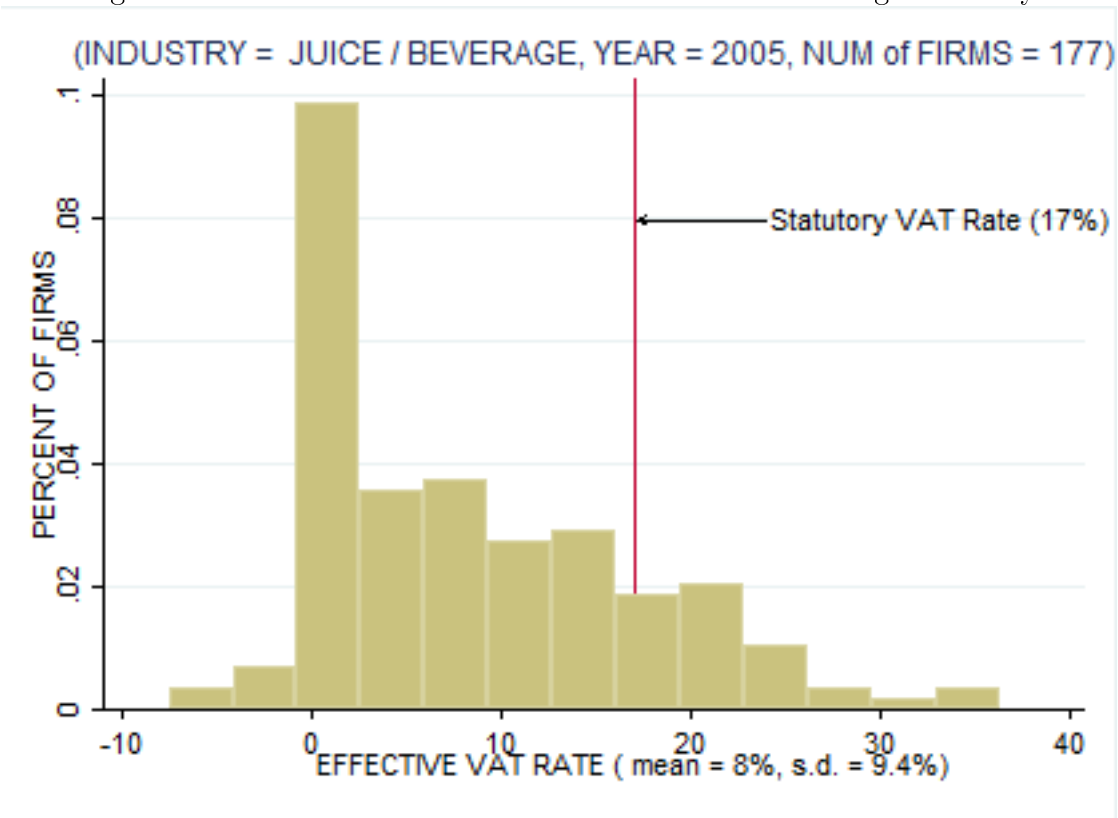


Figure 4.1 depicts the VAT rate dispersion within a four-digit industry (vegetable/fruit juice or beverages, industry code=1533) in the year 2005. The standard deviation of VAT rate in this industry is the median among all four-digit industries. Compared with Figure 1.1, we do not see significant drop in VAT rate dispersion. This is consistent with the finding in Table 1: most of the dispersion comes from within the industry. Details on

[6]Refer to Chinese literature by Lian He Ke Ti Zu (2007), Tan and Liu (2007).

analysis of potential sources of the dispersion and related empirical facts could be found in Appendix A.1.

5 Theoretical Framework: Distortions and TFP

5.1 Model

This model, largely based on Hsieh and Klenow (2009), aims to map the distortions in the firm level into the TFP loss in the aggregate level of the manufacturing sector. In order to fulfil this goal, we need a model which is able to aggregate the firm production, based on the firm survey data in China, into the total GDP of the whole manufacturing sector.

There are many sub-sectors (such as tobacco production, communication equipment and other similar sectors) within the manufacturing sector, which will be indexed by subscript s in the model . In the firm survey data (population census), s is particularly referred to the four-digit (two-digit) level industry. The classification and code of four-digit (and two-digit) level industry are based on the standard of "GB/T 4754—2002" promulgated by the Bureau of Statistics of China in the year 2002. In each sector s, there are lots of production firms, which are indexed by subscript i in the model.

There are, therefore, three tiers of firms in the manufacturing sector. 1. Firm at the bottom, indexed by si, produces Y_{si} using capital and labor as inputs; 2. Firm in the middle, indexed by s, produces Y_s with M_s differentiated inputs $\{Y_{si}\}_{i=1}^{M_s}$. 3. Firm at the top produces final output Y, using $\{Y_s\}_{s=1}^{S}$ as intermediate inputs. The firm at the bottom tier is real, and their data can be found in the firm survey. The middle and top tiers of firms are "dummy" and simply used to aggregate with certain technologies the products of the bottom firms.

The final output firm uses a Cobb-Douglas production technology:

$$Y = \prod_{s=1}^{S} Y_s^{\theta_s}, \text{where} \sum_{s=1}^{S} \theta_s = 1 \tag{5.1}$$

where Y is the total output in manufacturing sector, and Y_s is the intermediate input bought from sector s. HK(09) assumes a perfect competitive final output market. So the final good producer's objective is to maximize the profit by choosing $\{Y_s\}_{s=1}^{S}$ and taking prices $P, \{P_s\}_{s=1}^{S}$ as given.

$$\max_{\{Y_s\}} P \prod_{s=1}^{S} Y_s^{\theta_s} - \sum_{s=1}^{S} P_s Y_s \tag{5.2}$$

The profit maximization problem of the final-output firm yields the following FOCs:

$$P_s Y_s = \theta_s PY \tag{5.3}$$

Hereafter, we normalize P to 1 by taking final output as numeraire.

Intermediate firm-s produces Y_s with differentiated inputs Y_{si} by CES technology (Dixit-Stiglitz aggregator):

$$Y_s = \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{5.4}$$

where σ is the elasticity of substitution between different inputs. Here, the CES technology is introduced to accommodate the monopolistic competition in each four-digit sector s. The price mark-up for the firms in each sector is governed by σ.

The differentiated product Y_{si} is produced by the monopolistic competitive firm with a Cobb-Douglas technology of firm TFP A_{si}, physical capital K_{si}, and human capital H_{si}.

$$Y_{si} = A_{si} K_{si}^{\alpha_s} H_{si}^{1-\alpha_s} \tag{5.5}$$

where $H_{si} = h_{si} L_{si}$, and h_{si} is the firm specific average human capital level, L_{si} is the raw labour input.

Assume that R is the market rental price of physical capital, and w is the wage rate of raw labour. Both R and w are exogeneously given, so the model can be justifiably regarded as a partial equilibrium model.

There are two salient features of this model. One is the firm-specific exogeneous distortions in output ($\tau_{Y,i}$), capital ($\tilde{\tau}_{K_s i}$) and in labour ($\tilde{\tau}_{H_s i}$). The objective of firm si is to maximize the profit by choosing P_{si}, Y_{si}, and taking factor prices, distortions, and output demand curve as given.

$$\max_{\{P_{si},\, Y_{si}\}} \; (1-\tilde{\tau}_{Y_s i}) P_{si} Y_{si} - R(1+\tilde{\tau}_{K_s i}) K_{si} + w(1+\tilde{\tau}_{H_s i}) H_{si} \tag{5.6}$$

subject to the downward sloping (inverse) demand curve

$$P_{si} = P_s \left(\frac{Y_s}{Y_{si}} \right)^{1/\sigma} \tag{5.7}$$

where $\tilde{\tau}_{Y_s i}$=$1-\tau_{Y,i}-\tau_{si}^{VA}$, where τ_{si}^{VA} is firm specific VAT rate and $\tau_{Y,i}$ is other firm-specific distortions in output market.

For narrative purposes, we define the distortion-adjusted factor prices as below.

$$R_{si} = R(1+\tau_{K_s i}) \tag{5.8}$$

$$w_{si} = w(1+\tau_{H_s i}) \tag{5.9}$$

where $1+\tau_{Ksi}=\frac{1+\tilde{\tau}_{Ksi}}{1-\tilde{\tau}_{Ysi}}, 1+\tau_{Hsi}=\frac{1+\tilde{\tau}_{Hsi}}{1-\tilde{\tau}_{Ysi}}$. By their definition, R_{si} and w_{si} can be viewed as the firm-specific factor prices, which are taken as given by profit-maximizing firm si. Now the firm si's problem can be re-written as

$$\max \quad P_{si}A_{si}K_{si}^{\alpha_s}H_{si}^{1-\alpha_s} - R_{si}K_{si} - w_{si}H_{si} \tag{5.10}$$

which looks like the normal profit maximization problem of a firm without distortions.

From the firm's profit maximization conditions, we can express the output price for firm i in sector-s as:

$$P_{si} = \frac{\sigma}{\sigma-1}\frac{1}{A_{si}}\left(\frac{R_{si}}{\alpha_s}\right)^{\alpha_s}\left(\frac{w_{si}}{\beta_s}\right)^{1-\alpha_s} \tag{5.11}$$

where $\frac{1}{A_{si}}\left(\frac{R_{si}}{\alpha_s}\right)^{\alpha_s}\left(\frac{w_{si}}{\beta_s}\right)^{1-\alpha_s}$ is the marginal cost of producing Y_{si}, and $\frac{\sigma}{\sigma-1}$ is the mark-up.

In reality, the distortion-adjusted factor prices, R_{si} and w_{si}, are *unobservable*. But in this model, they can be written as the function of *observable* resource allocations as below.

$$R_{si} = \alpha_s\frac{\sigma-1}{\sigma}\frac{P_{si}Y_{si}}{K_{si}} \tag{5.12}$$

$$w_{si} = (1-\alpha_s)\frac{\sigma-1}{\sigma}\frac{P_{si}Y_{si}}{H_{si}} \tag{5.13}$$

where $P_{si}Y_{si}$ and K_{si} respectively are the value-added and physical capital stock of firm si, which are provided in the firm survey data. Human capital H_{si} is not available in firm survey data. In HK(09), they just use the wage bill as proxy for human capital. In this paper, we will use the more direct measurement of H_{si} based on the connection between the wage bill in the firm survey and the related information revealed in the population census.

Below we give two expressions that will be used in the next section. Actually, one can get the same results from monopolistic competition with Dixit-Stiglitz aggregator.

First, we can also obtain the expenditure ratio of intermediate input Y_{si} in sector s as follows:

$$\frac{P_{si}Y_{si}}{P_sY_s} = \left(\frac{P_{si}}{P_s}\right)^{1-\sigma} \tag{5.14}$$

Second, the output price index in sector-s is:

$$P_s = \left(\sum_{i=1}^{M_s} P_{si}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \tag{5.15}$$

5.2 TFP and Production Efficiency

In this section, I move on to look at the relationship between distortions and the TFP.

First we need to distinguish two sorts of productivity measurements. One is *physical* productivity, also known as TFP. Another is *revenue* productivity. For firm si, the physical productivity is just A_{si}, and the revenue productivity can be defined as below.

$$TFPR_{si} = \frac{P_{si}Y_{si}}{K_{si}^{\alpha_s}H_{si}^{1-\alpha_s}} = P_{si}A_{si} \tag{5.16}$$

$TFPR_{si}$ can also be expressed in terms of distortion-adjusted factor prices as follows.

$$TFPR_{si} = \frac{\sigma}{\sigma-1}\left(\frac{R_{si}}{\alpha_s}\right)^{\alpha_s}\left(\frac{w_{si}}{1-\alpha_s}\right)^{\beta_s} \tag{5.17}$$

If there are no distortions, i.e., $\tilde{\tau}_{Y_si} = \tilde{\tau}_{K_si} = \tilde{\tau}_{H_si} = 0$ and $R_{si} = R, w_{si} = w$, then there would be no variation in revenue productivity $TFPR_{si}$ in every sector.

In parallel, we can also define the revenue productivity for sector s as:

$$TFPR_s = \frac{P_sY_s}{K_{si}^{\alpha_s}H_{si}^{1-\alpha_s}} \tag{5.18}$$

where $P_sY_s = \sum_i P_{si}Y_{si}$.

From the results obtained before, $TFPR_s$ can be written as:

$$TFPR_s = \frac{\sigma}{\sigma-1}\left(\sum_i \frac{R_{si}}{\alpha_s}\frac{K_{si}}{K_s}\right)^{\alpha_s}\left(\sum_i \frac{w_{si}}{1-\alpha_s}\frac{H_{si}}{H_s}\right)^{1-\alpha_s} \tag{5.19}$$

where $K_s = \sum_i K_{si}$, $H_s = \sum_i H_{si}$.

Define the TFP in sector s as below:

$$TFP_s = \frac{Y_s}{K_s^{\alpha_s}H_s^{1-\alpha_s}} \tag{5.20}$$

Then we have

$$\begin{aligned} TFP_s &= \frac{1}{P_s}\frac{P_sY_s}{K_s^{\alpha_s}H_s^{1-\alpha_s}} && (5.21)\\ &= \left[\sum_{i=1}^{M_s}\left(\frac{1}{P_{si}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} TFPR_s && (5.22)\\ &= \left[\sum_{i=1}^{M_s}\left(A_{si}\frac{TFPR_s}{TRPP_{si}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} && (5.23) \end{aligned}$$

If there are no distortions, i.e., $\tilde{\tau}_{Ysi} = \tilde{\tau}_{Ksi} = \tilde{\tau}_{Hsi} = 0$, we would have $TFPR_{si} =$

$TFPR_s$ for all i in each s. Then TFP_s would be simplified as:

$$TFP_s^e = \left[\sum_{i=1}^{M_s} (A_{si})^{\sigma-1}\right]^{\frac{1}{\sigma-1}} \tag{5.24}$$

We call TFP_s^e the efficient TFP in sector-s since it is the TFP when there are no distortions.

$$Y_s = TFP_s K_s^{\alpha_s} H_s^{1-\alpha_s} \tag{5.25}$$

Therefore, the TFP loss caused by distortions in sector-s can be measured by $\frac{TFP_s(\tau)}{TFP_s^e}$, where $TFP_s(\tau)$ is the TFP in sector s under a vector of distortion τ.

How to measure the contribution of VAT rate dispersion to TFP loss?

$$\text{Contribution of VAT Dispersion} = \frac{TFP(\tau'_{K_{si}}, \tau'_{H_{si}})}{TFP(\tau_{K_{si}}, \tau_{H_{si}})} - 1 \tag{5.26}$$

where τ_K and τ_H are vectors of distortions in K and H including the VAT rate. τ'_K and τ'_H exclude the VAT rate.

$$1 + \tau_{K_{si}} = \frac{1 + \tilde{\tau}_{K_{si}}}{1 - \tilde{\tau}_{Y_{si}}}, \qquad 1 + \tau_{H_{si}} = \frac{1 + \tilde{\tau}_{H_{si}}}{1 - \tilde{\tau}_{Y_{si}}} \tag{5.27}$$

Using $1 - \tilde{\tau}_{Y_{si}} = 1 - \tau_{Y_{si}} - \tau_{si}^{VAT} \approx (1 - \tau_{Y_{si}})(1 - \tau_{si}^{VAT})$, we have

$$\begin{aligned} 1 + \tau'_{K_{si}} &= \frac{1 + \tilde{\tau}_{K_{si}}}{1 - \tau_{Y_{si}}} \approx (1 + \tau_{K_{si}})(1 - \tau_{si}^{VAT}) \\ 1 + \tau'_{Hsi} &= \frac{1 + \tilde{\tau}_{H_{si}}}{1 - \tau_{Y_{si}}} \approx (1 + \tau_{H_{si}})(1 - \tau_{si}^{VAT}) \end{aligned} \tag{5.28}$$

6 Data

In this section, I briefly introduce to the reader two data sets I am going to use.

From the theoretical model in Section 4, we know that, for firm si, we need the value-added data for Y_{si}, physical capital stock data for K_{si}, human capital for H_{si}, wage compensation to human capital H_{si}. We also need to know the factor share α_s and output share θ_s for each sector s.

Our strategy is to collect Y_{si}, K_{si}, L_{si}, and wage compensation from the *firm survey data* and get H_{si} from *population census data*. Factor share α_s, and output share θ_s are calculated based on the *firm survey data*.

6.1 Firm Survey

The firm data set is the same as the one used by Hsieh and Klenow (2009). It is the Annual Survey of Industrial Production conducted by the Chinese National Bureau of Statistics, which includes all non-state firms with more than 5 million RMB Yuan in revenue (about $600,000) plus all state-owned firms. I use the data from the year 2000 to 2007. There are over 160,000 firms in each year. Each firm has a four-digit industry code, ownership, region code, wage payments, value-added, capital stock, and a number of employees.

I use total asset as the physical capital stock, while HK(09) use the book value of the fixed capital net of depreciation. The main discrepancy between the two are current assets and intangible assets, which are supposed to create values-added to firms as well and therefore should not be ruled out once they are available in the dataset. Otherwise, I would get incorrect TFP for each firm and thus inappropriate TFPR dispersion across firms. Of course, we would repeat our work in the robustness check by using HK(2009)'s measurement for physical capital stock.

The wage compensation includes: (1) payable wage; (2) unemployment insurance premium; (3) pension and medical insurance premium; (4) housing mutual fund; and (5) total welfare fees. The wage compensation as a share of value-added is only 34% on average, while it is about 50% in an input-output table. To tackle this problem, I use the same method as HK(2009) by adjusting the wage compensation with a constant factor (multiplied by 3.6 in our case) such that the average wage compensation share is approximately 50%.

In ensure data quality, I dropped three types of observations: (1) the observation with empty cell and (2) the observation with non-positive value for capital stock, labour input, value-added and wage compensation, and (3) the observation with value-added less than the wage (even without adjustment in wage, the value-added in many firms is also less than the wage compensation). In the end, we are left with number of firms ranging from approximately 66,000 to 168,000 in each year during the period from 2000 to 2007.

Table 2 briefly reports the summary statistics of several key variables in the firm survey from the year 2000 to 2007.

Table 2: Summary Statistics - Firm Survey (2000 - 2007)

6.2 Population Census

The population data is from the 1% Sampling Population Census in year 2005 conducted by the Chinese National Bureau of Statistics. There are 2,585,481 individuals in the data set. The variables I use in this paper are: educational attainment, completion of

Table 3: Summary Statistics - Population Census (2005)

schooling, main source of income, two-digit industry code, and the ownership of the firm where the individual is working. To check the robustness of the results of this paper in the later stage, I may still want to use other variables such as: year and month of birth, sexual orientation, two-digit address code, ethnicity, monthly income, and occupation.

I dropped the following observations: (1) those not working in the manufacturing sector; (2) those whose occupation is 'farmer' or non-firm employee; (3) those whose main source of income is not from labour and from assets. After this filtering process we were finally left with about 200,000 observations.

The population census data and the firm survey data can be matched in the year 2005 by the industry-provincial-ownership characteristics. There are thirty-nine two-digit manufacturing industries, thirty-one provinces, and two types of ownership (state-owned and non-state-owned). The match between the two data sets allows us to recover the human capital for each firm in the firm survey data, as will be shown in the next section.

Table 3 briefly reports the key variables in the population census.

Another important issue is the measurement of human capital. In Hsieh and Klenow (2009), they use the wage bill of each firm to proxy its human capital. In this paper, I attempt to improve the human capital measurement by incorporating information about wages, years of schooling and other relevant characteristics of the labour force in the 1% population sampling survey. Details of this work are discussed in the Appendix.

7 Main Results

To numerically implement the computation of contribution of distortions to TFP loss, in the first instance I need to calibrate several parameters.

Following HK(2009), I set $\sigma = 3$. HK(2009) calibrate α_s, capital share in China, to the corresponding number in the U.S. In this paper, I calibrate α_s to the firm survey data in China, for which I have two reasons. First, unlike HK(2009), I do not compare the distortions between China and the U.S., so I do not need the U.S. as a benchmark. Second, in theory, the appropriate technology in China should be different from that in the U.S. due to different endowments (say, the proportion of skilled labour) in these two countries. It is hard to believe that China's capital share would be the same as that of U.S.

Regarding R and w, I calibrate them, based on the firm survey data, respectively to the average rental price of capital and wage rate in the whole manufactural sector.

Then, following the steps below, I can calculate the contribution of distortions to TFP loss.

Step 1: Based on Eq. (5.12) and Eq. (5.13), R_{si} and w_{si} can be calculated from the value-added $P_{si}Y_{si}$, physical capital stock K_{si}, and human capital stock H_{si}. $P_{si}Y_{si}$ and K_{si} are available from the firm survey data. H_{si} is the human capital measurement, either by wage bill as proxy, or by the approach proposed by this paper (I will discuss this point in detail presently)..

Step 2: Given the results of R_{si} and w_{si} obtained from step (1), I can back out τ_{K_si} and τ_{H_si} based on Eq. (5.8) and Eq. (5.9).

Step 3: Calculate $TFPR_{si}$ based on Eq. (5.17).

Step 4: Calculate $TFPR_s$ based on Eq. (5.19).

Step 5: Calculate TFP_s based on Eq. (5.21).

Step 6: Following HK(2009), A_{si} can be calculated from the following equation: $A_{si} = \frac{(P_{si}Y_{si})^{\frac{\sigma}{\sigma-1}}}{K_{si}^{\alpha_s} H_{si}^{1-\alpha_s}}$.

Step 7: Calculate TFP_s^e based on Eq. (5.24).

Step 8: Calculate the *Contribution of VAT dispersion* from Eq. (5.26).

7.1 VAT Rate Dispersion and TFP/GDP Loss

This section reports the most important result in this paper – the efficiency loss caused by the VAT rate dispersion across manufacturing firms in China, which is calculated based on the counterfactual analysis. I will do it with two sets of distortions in the output market in our exercises: (1) All the distortions $\tilde{\tau}_{Y_{si}} (= \tau_{Y_{si}} + \tau_{si}^{VAT})$; (2) Only $\tau_{Y_{si}}$, the distortions excluding firm specific VAT rate $\tau_{Y_{si}}$. Here I assume τ_{si}^{VAT} is independent of $\tau_{Y_{si}}$, $\tilde{\tau}_{K_{si}}$, $\tilde{\tau}_{H_{si}}$, and A_{si}. This assumption justifies that the distribution of other distortions ($\tau_{Y_{si}}$, $\tilde{\tau}_{K_{si}}$, and $\tilde{\tau}_{H_{si}}$) would not change when I eliminate the VAT rate dispersion (τ_{si}^{VAT}).

In both counterfactual exercises, I regard TFP^e, the efficient TFP in a world without any distortions, as the benchmark.

Table 4: Dispersions in VAT and TFP/GDP Loss

I first calculate the TFP loss caused by all the distortions $\tilde{\tau}_{Y_{si}}$, which is measured by $1 - TFP(\tau'_{K_{si}}, \tau'_{H_{si}})/TFP^e$, where $\tau'_{K_{si}}, \tau'_{H_{si}}$ respectively are the distortions in K and H normalized by $\tilde{\tau}_{Y_{si}}$ based on Eq.(5.27). The results every year are reported in row (1) (TFP Loss) of Table 4.

Then I calculate the TFP loss caused by all the distortions $\tilde{\tau}_{Y_{si}}$, measured by $1 - TFP(\tau_{K_{si}}, \tau_{H_{si}})/TFP^e$, where $\tau_{K_{si}}, \tau_{H_{si}}$ are the distortions in K and H normalized by $\tau_{Y_{si}}$ based on Eq.(5.28). The results are reported in row (2) (TFP Loss Net of VAT).

The efficiency loss caused by VAT rate dispersion is measured by $\frac{TFP(\tau_{K_{si}}, \tau_{H_{si}})}{TFP(\tau'_{K_{si}}, \tau'_{H_{si}})} - 1$, which can be interpreted as the percentage GDP/TFP gain from removing the VAT rate dispersion (τ_{si}^{VAT}) compared with the current actual GDP/TFP ($TFP(\tau'_{K_{si}}, \tau'_{H_{si}})$). This

exercise is close to the revenue-neutral tax reform since it only eliminates the spread between the firm-specific VAT rate and the economy-wide average VAT rate, with the tax rate for the high-tax firm being lowered and for the low-tax firm being raised. The results are in row (3) (Contribution of VAT). The mean of the effective VAT rate within industries are reported in row (4).

From numbers in row (3) I can see that the contribution of VAT rate dispersion to GDP (or TFP) loss is characterized by its significant magnitude and high volatility. The percentage of GDP (or TFP) loss ranges from 5% to 16%, while the average effective tax rate is largely unaffected.

7.2 Distribution of τ_K, τ_H, τ^{VAT}, and their correlation

In the counter-factual analysis I have done, I assume τ_{si}^{VAT} is independent of $\tau_{Y_{si}}$, $\tilde{\tau}_{K_{si}}$, $\tilde{\tau}_{H_{si}}$, and A_{si}. However, how plausible is this assumption? Also, to what extent will the conclusion will be changed if this assumption is violated?

Table 5 shows the correlation coefficients between five variables in the model. The numbers in the first column suggest that the correlation between the VAT rate, other distortions, and productivity $\ln A_{si}$ are very low.

Table 5: Correlation between Distortions and Productivity

If $\ln(1-\tau_{si}^{VAT})$, $\ln(1+\tau_{K,si}^{'})$, $\ln(1+\tau_{H,si}^{'})$, $\ln A_{si}$ satisfy the normal distribution, then we have the following expression for TFP in a four-digit industry s.

$$\ln TFP_s = \frac{1}{\sigma-1} log\left(\sum_{i=1}^{M_s} A_{si}^{\sigma-1}\right) - \frac{\sigma-1}{2} var\left(\ln TFPR_{si}\right)$$

which implies the TFP loss is caused by the dispersion in $TFPR_{si}$ across firms, where

$$\begin{aligned} var\left(\ln TFPR_{si}\right) &= var\left(\ln(1-\tau_{si}^{VAT}) + \ln TFPR_{si}^{KH}\right) && (7.1) \\ &= var\left(\ln(1-\tau_{si}^{VAT})\right) + var(\log TFPR_{si}^{KH}) \\ &\quad +2cov(\ln(1-\tau_{si}^{VAT}), \ln TFPR_{si}^{KH}) && (7.2) \end{aligned}$$

where $\ln TFPR_{si}^{KH} = \alpha_s \log(1+\tau_{K,si}^{'}) + (1-\alpha_s)\log(1+\tau_{H,si}^{'})$, and Table 5 shows the correlation between $\ln(1-\tau_{si}^{VAT})$ and $\ln TFPR_{si}^{KH}$ is equal to 0.0009, very close to zero.

In our previous analysis, we assume τ_{si}^{VAT} is independent of $\tau_{Y_{si}}$, $\tilde{\tau}_{K_{si}}$, $\tilde{\tau}_{H_{si}}$, and A_{si}. Now we know that the error we made in calculating the percentage contribution of VAT rate dispersion to TFP can be captured approximately by $\frac{\sigma-1}{2} \times 2cov(\log(1-\tau_{si}^{VAT}), \log TFPR_{si}^{KH})$. Given $\sigma = 3$, we have $\frac{\sigma-1}{2} = 1$.The number is equal to $2 \times 0.0009 \times 0.1451 \times 1.1036 \times 100\% = 0.0288\%$. So we underestimate the TFP loss caused by VAT rate dispersion in Table 4 only by a negligible magnitude of 0.0288%.

8 Robustness Check

8.1 Export Refund

To see to what extent the export tax refund leads to the VAT rate dispersion and efficiency loss, we can do the robustness check by recovering the counter-factual VAT rate for each firm by removing the export refund in the same way as we did in Table 19. The results are reported in Table 6, where row (1) is the contribution of VAT rate dispersion to GDP/TFP loss after removing the export refund, and row (2) is the benchmark case with export refund. Compared to row (2), the numbers in row (1) are not significantly and systematically larger or smaller, suggesting a minor role of export tax refund in TFP loss.

Table 6: TFP Loss with VAT Tax Refund Excluded

8.2 Time Average

To see what percentage of the efficiency loss caused by VAT rate dispersion can be attributed to firm-fixed effect and what percentage to time-idiosyncratic shock, we take the time average of VAT rate over the firms in six sub-samples. To do this, we first need to link the firm over years. However, since the firm ID changed in the year 2004, we can only link the firms respectively in two periods: 2000 to 2003, and 2005 to 2007. In each period, we consider several sub-samples of firms which stay for certain years.

For every firm in each sample, we take the year average of the VAT rate. In Sample 1, the year average is taken over four years for all firms in the sample. In Samples 2 and 5, the year average is taken over three years for all firms in the sample. In Samples 3 and 6, the year average will be actually taken over only two years for some firms. In Samples 4 and 7, the year average is just the actual VAT rate in a year within the period.

Using the same method to calculate the "Contribution of VAT" to TFP/GDP loss as in Section 7, for each sub-sample of firms I report in Table 7 and 8 the "Contribution of VAT" of actual VAT rate in each year and the year-average over the period.

Table 7: Time Average of VAT Rate (2000-2003)

Table 8: Time Average of VAT Rate (2005-2007)

For firms staying for the whole period (Samples 1 or 4), by comparing the numbers in different columns, the "Contribution of VAT" of year-average (1.5% in 2000 to 2003, 4.0 in 2005 to 2007) does not differ too much from that of each year, except in the

years 2002 and 2007. However, the years 2002 (4.9%) and 2007 (12.6%) themselves differ significantly from other years in each period. This implies that the year-average does not reduce the contribution of VAT, and the firm-fixed effect plays a major role in the VAT rate dispersion across firms.

Then we can make a comparison between different rows along the same column. We can clearly see that the "Contribution of VAT" is increasing in the sample size (Sample 4 > Sample 3 > Sample 2 > Sample 1, and Sample 7 > Sample 6 > Sample 5). This suggests that the entry and exit of firms brings about additional contribution to TFP loss. This is mainly because the entry-exit firms are more distorted in capital and labour market and thus reduce the actual $TFP(\tau'_{K_{si}}, \tau'_{H_{si}})$ and increase the contribution $TFP(\tau_{K_{si}}, \tau_{H_{si}})/TFP(\tau'_{K_{si}}, \tau'_{H_{si}}) - 1$, rather than because the VAT rates are more dispersed over these firms and reduce the $TFP(\tau_{K_{si}}, \tau_{H_{si}})$.

Why the years 2002 and 2007 are significantly different is a question that remains in need of an answer. One possible explanation is that the political cycles and the macroeconomic administration measures in these particular years, such as credit control, resulted in resource reallocations and distortions and thus reduced the productivity.

This could be an interesting direction for future research.

8.3 Data Quality

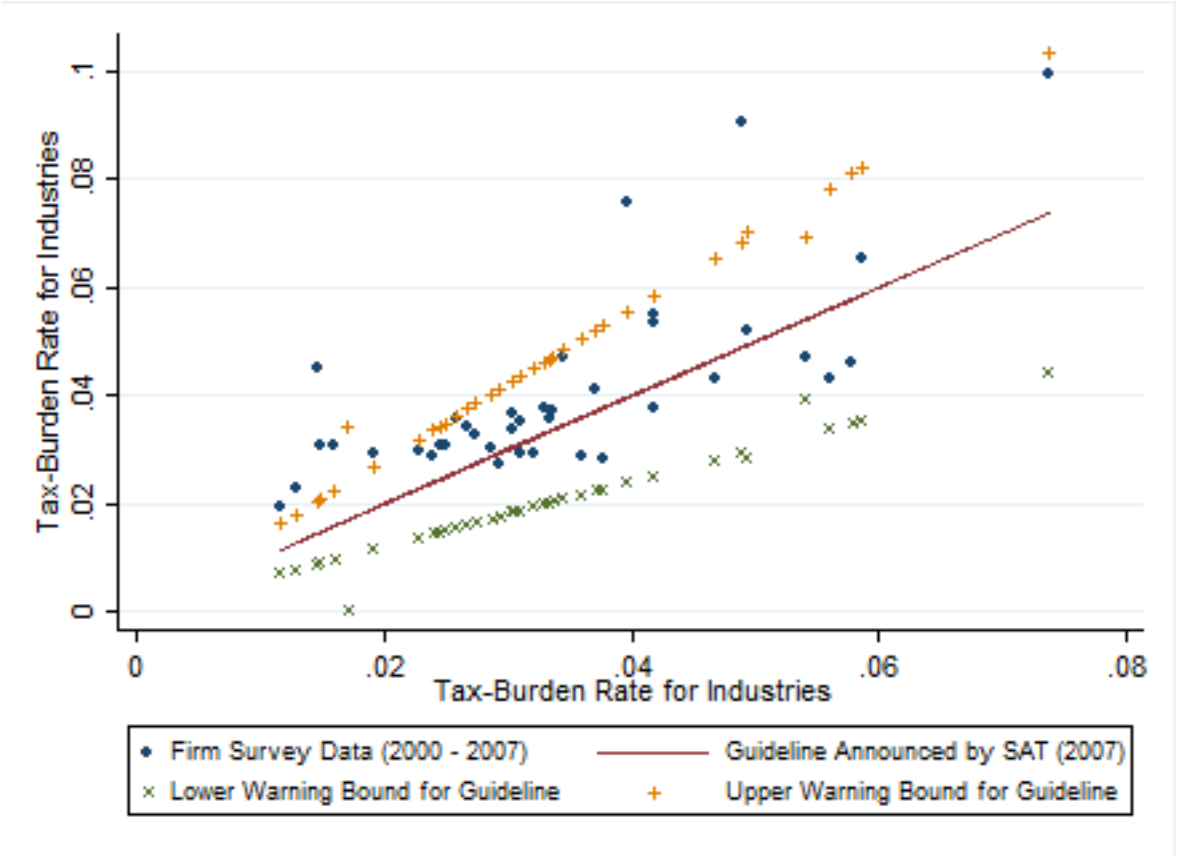
Another general and significant concern regarding the results in this paper is that of data quality, even though most of the tax experts and officials at the State Administration of Taxation assert how strictly the VAT is administered and enforced in China. The taxation bureau can locate any two VAT invoices for a specific transaction using an electronic network system. In this section, I still take this concern regarding data quality into serious consideration, by double-checking the data both from the Bureau of Statistics (the Firm Survey I use in this paper) and from the State Administration of Taxation.

Unfortunately, the firm level tax data is not available from the State Administration of Taxation. However, the State Administration of Taxation releases the *Tax Burden Ratio* for each two-digit industry and its lower bound as a warning signal regarding firms' tax evasion. *Tax Burden Ratio* is defined as $\frac{\text{Payable VAT}}{\text{Sales Value}}$. The State Administration of Taxation of China calculates the "Guideline" and "Lower Bound" of VAT burden ratio of each industry based on the firm level tax data over previous years. The "Guideline" is the average VAT burden ratio for firms within industry. The "Lower Bound" and "Upper Bound" is just the analogue to the confidence interval in statistics.

Figure 8.1 shows the "Guideline", "Lower Bound", and the "Upper Bound" from the State Administration of Taxation and the average VAT rate in the Firm Survey of each two-digit industry. Among 40 two-digit industries, 9 industries lie outside of the warning bound.

In the robustness check, firms in two-digit industry outside of the warning bounds are entirely removed from the sample.

Figure 8.1: Data Quality Check



Several conclusions can be drawn from Table 9 which reports the results after dropping the "badly" inconsistent industries. First, the contribution of VAT dispersion to GDP (or TFP loss) does not change systematically. That can be seen by comparing row (1) and row (2). Second, the average of the effective VAT rate in the remaining sample is almost unchanged. Third, the TFP loss caused by other distortions systematically decreases for all years. This implies that the firms in "badly" inconsistent industries also suffer from more distortions in capital and labour markets.

Table 9: Data Quality

8.4 Outlier

Another potential problem with the firm survey data is that outliers could exaggerate the contribution of VAT rate dispersion to TFP loss if a small number of firms extremely over-report or under-report their VAT liability and drive up the dispersion in the effective VAT rate. To avoid the contamination by outliers, I can drop $2\alpha\%$ firms from the whole

sample in each year if a firm's effective VAT rate is below $\alpha - th$ percentile or above $(100 - \alpha) - th$ percentile in that year.

Table 10 reports the results after dropping 0.1% of firms ($\alpha = 0.05$). I can see that the results are not affected, except in the years 2006 and 2007, where the contribution of VAT dispersion to TFP loss significantly decreases to the average level rather than much below the average.

Overall, the greater α is, the less dispersed is the effective VAT rate, and the contribution of VAT dispersion to TFP loss may also decrease. But is $\alpha = 0.05$ is a reasonable setting? The answer is probably 'yes'. To observe this, I report in Table 10 the minimum and maximum of the effective VAT rate after dropping the outliers. The minimum is -30% and the maximum is 100%. According to the tax law and tax practice in China, these are not extraordinary numbers even in the absence of tax evasion. Therefore, there is no strong evidence that we still need to raise the value of α.

Table 10: TFP Loss with Outliers Excluded

8.5 Elasticity of Substitution

As we have shown in Section 1, greater elasticity of substitution would translate into higher efficiency loss. This is because with a greater elasticity of substitution, a low-tax firm could steal more business from a high-tax firm and this could lead to more serious resource misallocation.

As with Hsieh and Klenow (2009), we also do the robustness check with greater elasticity by setting $\sigma = 5$ and report the result in row (1) of Table 11. Similar to Hsieh and Klenow (2009), the TFP loss of all distortions is aggravated massively, and the contribution of VAT is also enormously amplified compared to benchmark results in row (2) with $\sigma = 3$. This suggests that the results are sensitive to the parameter value of σ. More work needs to be done in estimating σ.

Table 11: TFP Loss with $\sigma = 5$

8.6 VAT Reform

Could the results be affected by the VAT reform in north-east provinces in 2004, as mentioned in Section **??** and suggested by Table 17? To answer this question, I dropped all firms in eight industries as well as the three provinces involved in the reform. The results are reported in row (1) of Table 12 and show that the contribution of VAT to TFP loss does not change very much compared to the benchmark results in row (2), except

where it decreases significantly in year 2007. In other years, it may increase or decrease a little bit, without a systematic direction of movement.

Table 12: TFP Loss with VAT Reform Provinces and Industries Excluded

9 Conclusion

Among all possible distortions that lead to misallocation and produce efficiency loss, although the effective VAT rate might not be the most salient one, it is a measurable one. This allows us to study the impact of dispersion in the effective VAT rate on the aggregate production efficiency. Based on the structural model and approach initiated by Hsieh and Klenow (2009), I conduct a revenue-neutral counter-factual analysis in this paper by removing the VAT rate dispersion to calculate the potential TFP/GDP gain. The results show that the TFP/GDP gain is as large as an average order of 7.9%, and volatile too, ranging from 5% to 16% in the period from 2000 to 2007. This suggests that the VAT in China can lead to non-negligible distortions and misallocation.

Robustness checks suggest that the efficiency loss mainly comes from the time-persistent and firm-fixed factors rather than the time-varing shock, demonstrating that the data quality is not a problem, but revealing that the results are sensitive to the parameter value of elasticity of substitution. The correlations are close to zero between the firm's VAT rate and some of the firm's characteristics, including firm size, TFP, distortions in capital and labour. This suggests that the distortion of VAT is very likely to be independent of other potential distortions in China.

One caveat should be noted again in the end. The effective VAT rate dispersion could be caused either by intricate real-world tax codes or by poor tax administration. This paper is silent on the respective role of each. Further studies could be done by separating the contribution of the two to the production efficiency. An additional feasible way is to employ the firm level data in developed countries such as UK and France, and to set their effective VAT distribution as a benchmark in the counter-factual exercise, rather than using the dispersion-free distribution as the benchmark as I have done in this paper. This is because the effective VAT rate dispersion is due to two broad factors: tax law (e.g., the existence of a general and reduced rates, rebates for exports and zero-rated goods) and the implementation of the policy (e.g., tax evasion, corruption, etc.), the TFP efficiency loss calculated in this paper should be attributed to both factors. If one believes that the dispersion in the effective VAT rate in developed countries are much less likely to be subject to bad tax implementation than in China, then the counter-factual exercise in that way may be able to tell us the TFP loss only caused by the bad tax implementation, rather than by the various special tax treatment.

References

1. Bartelsman, E. J., Haltiwanger, J. C. and Scarpetta, S. (2009) "Cross-Country Differences in Productivity: The Role of Allocation and Selection", *NBER Working Paper* No. 15490.

2. Bulutoglu, K. (1995) "Presumptive Taxation", in *Tax Policy Handbook*, edited by Parthasarathi Shome, pp. 258-262. International Monetary Fund.

3. Cai, H. B. and Liu, Q. (2009) "Competition and Corporate Tax Avoidance: Evidence from Chinese Industrial Firms", *Economic Journal*, Royal Economic Society, vol. 119(537), pp. 764-795.

4. Caselli, F. (2005) "Accounting for Cross-Country Income Differences", In Handbook of Economic Growth, Vol. 1A, ed. Philippe Aghion and Steven N. Durlauf, pp. 679–741. New York: North-Holland.

5. Caselli, F. and Gennaioli, N. (2006) "Dynastic Management," *CEP Discussion Papers* dp0741, Centre for Economic Performance, LSE.

6. Chen, S. Xiaoguang. (2015a) "The Effect of a Fiscal Squeeze on Tax Enforcement: Evidence from a Natural Experiment in China", *Working Paper*, London School of Economics.

7. Chen, S. X. G. (2015b) "Political Determinants of Selective Tax Enforcement: Evidence from China", *Working paper*. UK: *London School of Economics*.

8. Diamond, P. A. and James A. M. (1971) "Optimal Taxation and Public Production I: Production Efficiency", *American Economic Review,* 61(1), pp. 8-27.

9. Dixit, A. K. and Stiglitz, J. E. (1977) "Monopolistic Competition and Optimum Product Diversity", American Economic Review, 67(3), pp. 297-308.

10. Feldstein, M. S. (1978) "The Welfare Cost of Capital Income Taxation", *Journal of Political Economy*, vol. 86(2), pp. S29-S51.

11. Feldstein, M. S. (1999) "Tax Avoidance And The Deadweight Loss Of The Income Tax", *The Review of Economics and Statistics*, 81(4), pp. 674-680.

12. Gordon, R. and Wei, L. (2009). "Tax structures in developing countries: Many puzzles and a possible explanation", *Journal of Public Economics*, Elsevier, vol. 93(7-8), pp. 855-866.

13. Guner, N., Ventura, G. and Xu. Y. (2008) "Macroeconomic Implications of Size-Dependent Policies", *Review of Economic Dynamics*, vol. 11(4), pp. 721-744.

14. Hsieh, C. T. and Klenow, P. J. (2009) "Misallocation and Manufacturing TFP in China and India", *The Quarterly Journal of Economics*, vol. 124(4), pp. 1403-1448.

15. Hsieh, C. T. and Klenow, P. J. (2010) "Development Accounting", *American Economic Journal: Macroeconomics, American Economic Association*, vol. 2(1), pp. 207-23, January.

16. Harberger, A. C. 1964(a) "The measurement of waste", *American Economic Review*, 54 (3), pp. 58-76.

17. Harberger, A. C. 1964(b). "Taxation, resource allocation, and welfare", in John F. Due ed., *The role of direct and indirect taxes in the federal revenue system* (Princeton, NJ: Princeton University Press), pp. 25-70.

18. Hines, J. R. Jr. (1999) "Three Sides of Harberger Triangles", *Journal of Economic Perspectives*, 13(2), pp. 167-188.

19. Jones, C. (2011) "Intermediate Goods and Weak Links in the Theory of Economic Development", *American Economic Journal: Macroeconomics*, 3 (2), pp. 1-28.

20. Krueger, A. B., Summers, L. H. (1968) "Reflections on the Inter-Industry Wage Structure", *NBER Working Paper*, No. 1986.

21. Lian He Ke Ti Zu. (2007) "Shenzhen Shi Zhong Dian Shui Yuan Fen Xi Ji Jian Yi (in Chinese)", *Qi Ye Jin Ji*, 11.

22. Mirrlees, J., Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G. and Poterba, J.. (2011) *Tax by Design: the Mirrlees Review*, Oxford University Press.

23. Moll, B. and Banerjee, A. (2010) "Why Does Misallocation Persist?" *American Economic Journal: Macroeconomics*, 2(1), pp. 189-206.

24. Restuccia, D. and Rogerson, R. (2008) "Policy Distortions and Aggregate Productivity with Heterogeneous Plants", *Review of Economic Dynamics*, vol. 11(4), pp. 707-720.

25. Stotsky, J. (1995) "Minimum Taxes", in *Tax Policy Handbook*, edited by Parthasarathi Shome, page 263-266. International Monetary Fund.

26. Tan, J. Y. and Liu, A. W.. (2007) "Yi Xin Xi Hua Wei Yi Tuo Shi Shi Shui Yuan Jin Xi Hua Guan Li De Si Kao (in Chinese)," *Guang Dong Di Fang Shui Wu Za Zhi*, 05.

27. Zee, H. H. (2005) "VAT Treatment of Financial Services: A Primer on Conceptual Issues and Country Practices", *Intertax*, 34(10), 458-474.

28. Wikipedia (2015) *Hukou system*. [Online] Available from: http://en.wikipedia.org/wiki/Hukou_system [Accessed: 1st June, 2015]

Appendix 1. Effective VAT Rate Dispersion – Potential Sources and Facts

A.1.1 Potential Sources of VAT Dispersion

Theoretically, Value Added Tax (VAT) avoids the cascade effect of sales tax by taxing only the value added at each stage of production. An ideal VAT system is supposed to apply a uniform tax rate at every stage along the value-added chain, concluding at the final consumption. Without doubt, the effective VAT rate for different taxpayers would be the same under a perfect VAT system. In practice, however, the effective VAT rate varies across firms due to a variety of reasons: (1) The breakdown of the value-added chain; (2) A non-uniform tax rate; (3) Issues relating to tax compliance; (4) Issues relating to tax administration and enforcement; (5) Tax reforms. To be consistent with the motivation of this paper, I focus primarily on the sources of the VAT rate dispersion within the same (four-digit) industry, as opposed to discussing the VAT rate differentials between firms in different industries.

The sources of effective VAT rate dispersion can be demonstrated by a credit-invoice computation method:

$$\tau = \frac{p_y q_y \tau_y - p_m q_m \tau_m}{p_y q_y - p_m q_m} = \left(1 - \frac{p_m q_m}{p_y q_y}\right)^{-1} (\tau_y - \tau_m) + \tau_m \tag{9.1}$$

where τ is the effective VAT rate of a firm. p_k, q_k, t_k $(k = y, m)$ respectively are price, quantity, and tax rate for goods k ($k = y$ for output, $k = m$ for intermediate input). $p_y q_y - p_m q_m$ is value-added. $p_y q_y \tau_y$ is the tax payment on sales, $p_m q_m \tau_m$ is the reclaimable tax on intermediate goods purchase. $p_y q_y \tau_y - p_m q_m \tau_m$ is payable VAT.

Theoretically, if $\tau_y = \tau_m = 17\%$ for all firms, then there would be no cross-firm dispersion in τ. But dispersion arises in two cases: (1) Even though all firms have the same τ_y and the same τ_m, while if $\tau_y \neq \tau_m$, then the disparity between firms in $(p_m q_m)/(p_y q_y)$ would imply the dispersion in τ. (2) If $\tau_y = \tau_m$ for each firm, but they vary across firms, then we still have dispersion in τ. These two cases may arise from several sources including tax law, tax compliance, tax administration and enforcement, and tax reform in China, and this is discussed in greater detail below.

Break-down of the Value-added Chain

If all the transactions are included in the VAT system and are applied with the same flat tax rate, there would be no tax rate dispersion across firms. But the breakdown of the value-added chain gives rise to VAT rate dispersions; below are two common cases.

First, service sectors are excluded from VAT system. In China, exemption of many service sectors from the VAT system does not allow the manufacturing firms that purchase the service from other firms to reclaim on its service input. In this case, $\tau_m = 0$ for the service intermediate input purchased from other firms. Even though all firms face the same, τ_y, the firm with smaller $(p_m q_m)/(p_y q_y)$ benefits from lower τ. A manufacturing firm, therefore, has an incentive to vertically integrate service firms in order to reduce the purchase of intermediate input $p_m q_m$. Different $(p_m q_m)/(p_y q_y)$ across firms, probably due to firms' heterogeneous capability of vertical integration, would result in tax rate dispersion.

Second, invoices are difficult to get from purchasing agricultural goods. In China, it is very common that manufacturing firms (say, tobacco factories) purchase the agricultural goods (say, tobacco leaves) as their raw material from farmers who in general are not able to provide the invoice. In this case, $\tau_m = 0$. By the same token, a firm will face a low τ if it sets up an agricultural farm of its own, or if it is able to purchase from large formal farms which can provide invoices for their agricultural products.

Non-uniform Tax Rate

There are three common cases of deviation from the standard VAT rate at 17%.

First, differential tax rates. In China, the VAT rate is 17% at the standard rate but only 13% for agricultural goods. In this case, $\tau_y > \tau_m$. A firm can lower τ by vertically integrating an agriculture farm as a supplier of intermediate input.

Second, export refund and zero rating. In this case, $\tau_y < \tau_m$. Suppose two firms produce the same product. Firm D produces solely for the domestic market, while Firm E exports to foreign markets. Firm E can get the VAT tax refund and enjoy a lower effective tax rate than Firm D.

Third, purchase of goods from small taxpayers. In China, a tax reclaim on a purchase from small taxpayers is lower than the standard rate, which implies $\tau_y > \tau_m$. In this case, a firm will have a higher tax rate if it has to purchase the intermediate input from small taxpayers.

Tax Compliance

Although the system of VAT is favourable to other systems because it limits the scope of tax evasion and fraud by tracing down the invoices in transaction, tax compliance is

still a big concern in practice. HMRC in the UK estimates that the "VAT gap" was 11.5 billion, accounting for 14% of the potential revenue yield, from 2009 to 2010; that is the difference between tax actually collected and the tax that would have been paid. (Mirrlees *et al.*, 2011).

Non-compliance can occur in the following ways:

First, firms do not record sales that ought to be taxable. A firm can open a direct sales store without issuing invoices to consumers, or launch a sales promotion just by bestowing goods as opposed to lowering the price. Thus τ_y could be different between firms.

Second, invoices for input purchases can be faked. A firm can buy agricultural goods from non-VAT taxpayers. So τ_m could vary across firms.

Third, firms can claim that sales are lower-rated or zero-rated. A firm can fake export invoices. τ_m could vary across firms.

In addition, China still has a non-negligible number of SOE, and these may be more compliant when paying tax then Non-SOE.

Tax Administration and Enforcement

China is a huge country with substantial heterogeneity across geographical characteristics. VAT is collected by SAT (State Administration of Taxation). Under the hierarchical structure of SAT, regional differences in tax administration and enforcement can be enormous, at least for two reasons. First, the principal-agent relationship between taxation officials at different levels of SAT sets the asymmetric information problem at the central stage of tax collection. Second, taxation officials at the lower level of SAT, for instance at the county level, are very likely to be connected with the local government officials, and thus act in favour of local interests rather than in accordance with tax law. Regional VAT rates, therefore, are determined to a large extent by regional fixed factors such as the taxation capability, the agency cost, the administration cost, and the incentive of taxation enforcement.

In addition, the delay in making payment of payable VAT is also a common practice in China. It usually takes place when a firm has used up all their cash and has to make a request to the taxation office for an extension of time to make their tax payment. In this case, a firm's τ could vary over time, but its overall time average should remain stable.

Tax Reform

Tax reforms are generally launched initially only in some selected industries and regions, as opposed to implementing them all at once across all industries and in the whole country. This approach to reform may create tax rate dispersions across firms in different regions and industries. For example, VAT tax reform was initiated in eight industries in

the north-east provinces on July 1, 2004, with VAT being transformed from a product type to a consumption type, which led the firms in these industries in north-east provinces to experience a lower effective VAT tax rate than those in other regions. In July 2007, the reform was extended to another twenty-six cities in six central-region provinces. In this case, τ_m could be different for firms in different regions within the same industry.

A.1.2 Empirical Facts on VAT Dispersion

Although there are many potential sources of VAT rate dispersion, as set out above, the main factor is the regional variation, which accounts for about 60% of within-industry VAT rate dispersion. Observable tax compliance and tax planning behaviours, such as vertical integration, size and ownership of the firm, do also significantly affect the VAT rate, but with a notably smaller contribution to VAT rate dispersion.

The results also show that consumption type VAT reform reduced the VAT rate dispersion in reform industries, while the time-idiosyncratic shock, export refund, entry and exit of firms are not a main source of tax rate dispersion.

Tax Administration, Compliance, Ownership

To empirically investigate the sources of VAT rate dispersion, I run the following regression, in an attempt to measure the impact of regional fixed effect, ownership type of firm, and various other characteristics of a firm including its size, input-output structure, reliance on exports, organisation of firm, after controlling for four-digit industry and year fixed effects. The regression equation is as follows:

$$\tau_{it} = \alpha + \gamma_t + \eta_{c,s} + \gamma SOE_i + \eta Size_{it} + \lambda Organization_{it} + \delta Export_{it} + \xi_{it} \quad (9.2)$$

where γ_t and $\eta_{c,s}$ respectively are year fixed effects and county-(4-digit)industry fixed effect, SOE_i is dummy for ownership type ($SOE_i = 1$ for Non-SOE). $Size_{it}$ is measured by logarithm of firms' value-added, $Export_{it}$ by $\frac{Export}{Total\ sales}$. Organization of a firm is measured by two variables: (1) input-output structure$\frac{p_y q_y - p_m q_m}{p_y q_y}$, (2) number of establishments.

Results in Table 13 bear some implications. The R^2 changes from close to zero in column (1) to 0.06 in column (2). This suggests that about 6% of the cross-firm variation in the effective VAT rate can be explained by the variation across (4-digit) industries. By the same token, the R^2 in column (3) and (4) implies that the county fixed effect and the (4-digit) industries-county fixed effect have explanation power at an magnitude of about 11% and 44% respectively. The results in column (5) suggests that, even after controlling for some observable firms' characteristics that should affect the effective rate, there are still about 45% of the variation remains unexplained.

Table 13: Sources of VAT Rate Dispersions

Since $\tau = \frac{p_y q_y \tau_y - p_m q_m \tau_m}{p_y q_y - p_m q_m} = \frac{p_y q_y}{p_y q_y - p_m q_m}\tau_y - \frac{p_m q_m}{p_y q_y - p_m q_m}\tau_m$, we also want to know the sources of dispersion in τ_y and τ_m. Similarly, I regress τ_y and τ_m on the same regressors as Table 13, where τ_m is measured by $(Input\ VAT)/(Intermediate\ goods\ purchase)$, τ_y is measured by $(Output\ VAT)/(sales\ turnover)$. The results are reported in Table 14 and Table 15.

The R^2 in column (3) of table 14 is 0.53. This is in contrast to that in column (1) and (2), and implies that about 53% of the variation in the effective input VAT rate can be explained by persistent county fixed effects, while the industry fixed effects have no impact at all. This might be explained by problematic local VAT administration, particularly due to firms using fake invoice for VAT reclaim, as suggested by Chen (2015a, b). By comparison, the R^2 in column (3) of table 15 is only 4%. And it is as big as 0.83 in column (5), suggesting that the variation in the effective output VAT rate is probably explained by the statutory factors rather than by poor tax administration.

Table 14: Determinants of Input VAT Rate

From Table 15 we can deduce the following:

(1) The county fixed effect can account for approximately 70% of the dispersion in τ_y across firms within an industry.

(2) A firm's ownership, size, export refund, and number of establishments account for approximately 19% of τ_y.

Table 15: Determinants of Output VAT Rate

Tax Reform

An eyeball test from the summary statistics in Table 16 suggests the consumption-type VAT reform in 2004 does lead to some change in VAT rate dispersion for the north-east provinces, dropping from around 13% before reform to 11% of the post-reform level.

To see the precise effect of tax reform on VAT rate, I employ the triple difference approach (DDD) in identification, by looking into the variation over periods (before and after reform), across regions (three north-east provinces and others), and across industries (industries under reform and others). I respectively regress the VAT rate of firm i in year t (τ_{it}) and standard deviation of τ_{it} within an industry-region-period cell ($Std(\tau_{it})$) on three dummies and their cross-terms as follows:

Table 16: VAT Rate and Reform - Summary Statistics

$$\begin{aligned}\tau_{it}(\text{or } Std(\tau_{it})) &= \alpha + \beta R_i^{NE} + \gamma I_i^s + \eta Y_t^{After} + \theta R_i^{NE} Y_t^{After} \\ &\quad + \delta I_i^s Y_t^{After} + \kappa R_i^{NE} I_i^s + \varphi R_i^{NE} I_i^s Y_t^{After} + \xi_{it} \end{aligned} \tag{9.3}$$

where the parameter of interest is φ.

The regression results are reported in Table 17. The reform in six industries in three north-east provinces reduces the effective VAT rate by 0.76% (insignificantly) and its standard deviation by 1.25% (significantly at 5% level). In Section 8, I will check whether the conclusion on TFP loss caused by VAT rate dispersion could be significantly changed by excluding reform region and reform industries from the sample.

Table 17: VAT Rate and Reform - Regression

Time-idiosyncratic Shock

One may explain part of the VAT rate dispersion as a result of taxpayer's time-idiosyncratic behaviour, such as the delay of VAT payment this year until next year because of cash constraints. Admittedly, this is not unusual in China's practice. If we suppose that this is an important factor affecting the tax rate variation, then the VAT rate dispersion across firms would be greatly reduced by taking the time average, however this is not supported empirically.

The VAT rate of each firm exhibits two salient features over time: (1) strong fixed effect, and (2) persistency.

Fixed effect is captured by τ_i, which is largely inherited from regional and firm's characteristics. Regression results in Table 13 confirm that about 60% of the fixed effect come from a firm's characteristics in terms of jurisdiction, ownership type, and industry.

Persistency is represented by $AR(1)$ parameter ρ. For firm i, serial correlation coefficient (conditional on i) is $corr_i(\tau_{it}, \tau_{it+1}) = \rho$. Table 18 reports the correlation coefficient from 2000 to 2007, showing ρ is around 0.6.

Table 18: Serial Correlation of VAT Rate over Years

Export Refund

One may argue that an export refund could be an important source of VAT rate dispersion, given the fact that firms in China are increasingly more reliant on international

markets. The data, however, does not agree with this argument. Table 19 demonstrates that the export tax refund does not significantly contribute to the VAT rate dispersion, even though it reduces the average VAT rate. The data only provides the VAT rate ex-post refund. I recover the VAT rate ex-ante refund by removing the contribution of tax refund based on the regression of the VAT rate on export.

Table 19: VAT Rate Dispersions and Export Refund

Entry and Exit of Firm

Quite commonly, the entry and exit of a firm produces massive extensive margins in addition to the existing firms. To check the impact of entry and exit on an average VAT rate dispersion, I categorize the firms into two groups: (1) firms staying in the sample for at least the period from 2000 to 2003 (or at least 2005 to 2007); (2) firms staying for less than three years. Table 20 reports the difference in VAT rate and within-group standard deviation between Group 1 and Group 2. Group 2 has a significantly lower tax rate but higher dispersion by 1.02%, which is not considered to be a striking difference.

Table 20: Effect of Entry and Exit of firms on VAT Rate

Correlation with Size of Firm

Regression in Table 13 suggests that the effective VAT rate would decrease by 1% if the size of the firm (measured by valued-added) was increased by 2.5 times. This effect is significant but not notably big. In this paper, I assume $\tau_{it}(\omega)$ is independent of other characteristics of a firm, such as TFP or distortion in labour and capital, that may affect firms' decisions regarding their size measured by value-added, labour input, and capital input.

This assumption is based on three justifications, as follows. (1) Suppose the negative regression coefficient represents the causal effect from firm size to VAT rate. In a dynamic setting, a firm may take this effect into account in its expanding plan. In the short run or a static setting, however, the effect would be hardly relevant to the firm's production decision since even the rapid expansion of the firm in the short-term would only yield a negligible drop in the VAT rate. (2) The regression coefficient may likely reflect the reverse causality from tax rate to firm size. In this case, the independence between the VAT rate and the firm's size-determining characteristics could still lead to the negative regression coefficient under the rationale of the misallocation theory: a low-tax firm becomes larger by stealing business from high-tax rivals. (3) Very importantly, Table 21

demonstrates that the correlation between VAT rate and firm size, measured in terms of several variables, is quite small.

Table 21: Correlation between VAT Rate and Firm Size[a]

Appendix 2. Measurement of Human Capital

Human capital measurement is vital to distortion issues, since the distorted wage is measured in the following way:

$$w_i = w(1+\tau_i) = \frac{\sigma-1}{\sigma}\frac{(1-\alpha)P_iY_i}{h_iL_i} \tag{9.4}$$

Obviously, bad measurement for h_i will be attributed to distortion τ_i

It should be noted that the term "human capital" used in this paper is a misnomer for "Marginal Product Value of Labour". Hereby I call it "human capital" just for narrative purpose. In this section, I study how to calculate the marginal product value of labour, rather than the narrowly defined human capital which is usually viewed as the returns to schooling, training, and health. To do this, we need to know the "market return" to workers with certain characteristics, which include the worker's characteristics, including years of schooling, age, sex, and occupation; and the working firm's characteristics such as ownership type, location, and industry. In labour economics, the wage regression is a very complicated question. It may suffer serious problems such as ability bias, selection bias, and many others. Moreover, quite a lot of factors that contribute to the wage compensation, such as worker-firm matching quality, externality of human capital (peer effects), are unobservable to econometricians. In this paper, I leave aside these micro-labour concerns, due to the lack of high quality data and because of the macroeconomic nature of our original question.

Hsieh and Klenow (2009) use the firm's wage bill to proxy human capital and point out that the wage bill in China's firm survey might be problematic.

In the following sections, I attempt to sort out this problem by proposing a method for human capital measurement. I first compare the results, at the "industry-province-ownership cell" level rather than the firm level, by using the wage bill as proxy for human capital (as per HK(2009)'s approach) and by using the hedonic wage as human capital. We will observe that Hsieh and Klenow (2009) may understate the contribution of labour distortion to TFP loss. After realising the pitfalls regarding using the wage bill as proxy for human capital, I propose a method, based on the hedonic wage, to calculate the human capital (marginal product value of labour) for each firm in the firm survey data.

A.2.1 Why Do We Need the Hedonic Wage?

To examine the potential risk of using the wage bill as proxy for human capital, we need to compare the results between the case that uses the wage bill as proxy for human capital and the case that uses other measurements for marginal product value of labour.

There are two natural candidates for the measurements of marginal product value of labour: (1) Mincerian type human capital; (2) The hedonic wage. The former is widely used in the development accounting exercise (Klenow and Rodrigue Clare, 1997; Caselli, 2005).

From the theoretical model in section 5, we can observe that the accuracy of measurement in labour distortion relies heavily on how correctly we are aware of the human capital of labour in each industry. The measurement of human capital by wage bill and by Mincerian human capital (return to schooling) are at two ends of the spectrum. Using the wage bill as proxy for human capital implies that the observed wage dispersions across firms are caused exclusively by variations in human capital rather than by distortions in the labour market, leading to a smaller contribution of labour market distortions to TFP loss. At the opposite end of the spectrum, measuring human capital simply with return to years of schooling demonstrates that the marginal value of labour depends simply on the years of schooling, and dispersions in wage net of the dispersions in return to years of schooling would potentially cause market distortions that we cannot observe.

In this paper, I take the middle field between the two extreme cases. It is well known that, in real world, there are actually enormous wage dispersions across industries, even after controlling for the education level and other characteristics of workers (Krueger and Summers, 1986). The more reasonable way to measure the marginal product value of labour is to estimate the wage return to each characteristic of the worker and the corresponding job. This could be achieved by regression with the population census data. The standard wage regression would simply allow us to estimate the return to years of schooling and to each industry, as well as return to other characteristics of a worker. Actually, this approach is just an application of hedonic pricing to the wage determination in a competitive labour market. I presumably think, due to perfect market competition, that the hedonic wage fully reflects the marginal productivity of labour, and that the deviation of the actual wage from the hedonic wage is caused by various distortions in the labour market (both outside and inside the firm).

The hedonic wage can be calculated following the two steps below.

First, in light of the wage regression in labour economics literature, we can run the regression below in the first place:

$$\ln w_j^{p,05} = \alpha + X_j\beta + Z_j\gamma + \varepsilon_i \tag{9.5}$$

where $\ln w_j^{p,05}$ is a logarithm of wage of individual-j in *2005 population census*, the su-

perscript p of $\ln w_j^{p,05}$ implies that the data comes from the "*population* census", and the superscript 05 indicates the year "2005", X_j is a vector of individual-j's characteristics including years of schooling, age, age^2, sex, occupation, hukou[7], migrant worker. Z_j is a vector of characteristics of firm where individual-j works, including ownership of the firm, location, industry (two-digit). The fitted value for $\ln w_j^p$ can be viewed as the hedonic wage of individual-j.

Second, the hedonic wage in each province, each two-digit industry, and each ownership type can be calculated by taking the conditional average over the individual's $\ln w_j$.

$$\ln w_{p,i,n}^{p,05} = E(\ln w_j^{p,05} | \text{province } p,\ \text{industry } i, \text{ownership } n) \tag{9.6}$$

$\ln w_{p,i,n}^{p,05}$ is calculated in order to match the wage rates in the firm survey. I collapse the data in STATA by industry, province, and ownership type to get an average hedonic wage rate of each "firm" (cell) in 31 provinces, 39 industries, and 2 types of ownership. Since the cell is not rectangular, we have only 1,860 "firm-cells" rather than 2,418 (31×39×2) firms. On average, there are 115 (=214,318/1,860) individuals in each "firm". So the average hedonic wage of each firm is, though not perfect, still quite representative.

With two types of human capital measurements available, we can follow the seven steps described at the beginning of Section 6 and calculate the TFP loss caused by distortions in capital and labour. The results are reported in Table 22.

From Table 22, we can see that the distortion in human capital can lead to TFP loss by 13%, greater than the role of physical capital (11%). Due to correlations between the distortions, the three distortions combined reduce TFP by 19%, less than the sum of the individual contributions ($11\% + 13\% = 24\%$).

Table 22 also reports the TFP loss by using HK(2009) method (the wage bill as proxy for human capital). From Table 22 we can see HK(2009) under-estimate the TFP loss by approximately 5%. This is mainly because HK(2009) use the wage bill as proxy for human capital. The wage bill in the firm survey has larger dispersion than the true human capital and would attribute the distortions to the variations in human capital. The contribution of distortions to TFP loss, therefore, is understated.

From Table 22, we are convinced that it is better to use the hedonic wage as a mea-

[7]Hukou is a household registration record that officially identifies a person as a resident of an area and includes identifying information such as name, parents, spouse, and date of birth. Originally, Hukou as a family registering system was in existence in China as early as the Xia Dynasty (c. 2100 BC - 1600 BC). The current Hukou system was promulgated in 1958 by the Chinese government in order to control the movement of people between urban and rural areas. Individuals were broadly categorized as a "rural" or "urban" residents. People who worked outside their authorized domain or geographical area would not qualify for grain rations, employer-provided housing, or health care. There were controls over education, employment, marriage and so on. (Wikipedia, 2015)

surement for human capital. But up to now we have simply obtained the hedonic wage for each "industry-province-ownership cell". In order to calculate the effect of distortions at the firm level, we still need the human capital measurement for each firm. This is the task of the following section.

A.2.2 Hedonic Wage at the Firm Level

What is the hedonic wage rate for each firm in our sample? This is a difficult question since the firm survey only offers the wage bill data and does not say anything about the characteristics, such as years of schooling or age, of the employees. To cope with this question, I combine the firm survey and the population census in year 2005 and make a statistical inference in the following way.

Firstly, since from the last section we have also known the hedonic wage in each province, each two-digit industry, and each ownership type $\ln w^{p,05}_{p,i,n}$, to find the relationship between the hedonic wage and the wage rate reported in the firm survey, we can regress $\ln w^{p,05}_{p,i,n}$ on $\ln w^{f,05}_{p,i,n}$, where $w^{f,05}_{p,i,n}$ is the average rate in province p, industry i, and ownership n from the "*firm* survey".

$$\ln w^{p,05}_{p,i,n} = \kappa + \phi \ln w^{f,05}_{p,i,n} + \xi_{p,i} \tag{9.7}$$

based on which we can get the linear projection of the hedonic wage from the wage rate reported in the firm survey.

Second, I assume that relationship between $\ln w^{p,05}_{p,i,n}$and $\ln w^{f,05}_{p,i,n}$ remains stable across firms and over the firm survey sample periods. Making this assumption, I can calculate the hedonic wage $\hat{w}_{i,t}$ for firm i in year t based on $w^{f}_{i,t}$, the corresponding reported wage in firm survey, as follows:

$$\hat{w}_{i,t} = \exp(\hat{\kappa} + \hat{\phi} \ln w^{f}_{i,t}) \tag{9.8}$$

where $\hat{\kappa}$ and $\hat{\phi}$ are estimators of κ and ϕ respectively.

By running OLS on Eq. (9.7), I get $\hat{\kappa} = 8.22(0.032)$, $\hat{\phi} = 0.235(0.0077)$. (The numbers in the brackets are standard deviation) Also, $R^2 = 0.3344$.

With this human capital measurement for each firm, now we can turn to the most important work of this paper — calculating the TFP loss caused by tax distortions.

Table 1: VAT Rate Dispersions within (four-digit) Industries

year	Obs	All Firms[1]	Within Industries[2]	Contribution of Within[3]
2000	66,200	11.8%	10.5%	97.0%
2001	74,511	11.9%	10.7%	89.4%
2002	83,285	11.0%	10.2%	92.2%
2003	97,993	10.9%	10.1%	90.8%
2004	131,306	11.5%	10.6%	90.6%
2005	129,398	11.3%	10.3%	94.9%
2006	147,696	11.1%	10.2%	95.3%
2007	168,876	10.5%	10.0%	95.1%

Notes:

1. This column lists the standard deviation of the VAT rate across all firms in the Annual Survey of Industrial Production in each year.

2. This column lists the mean taken over 4-digit industries on the within-4-digit-industry standard deviation of VAT rate.

3. This column lists the contribution of within-industry VAT rate variation to the total variation. The results are obtained by variance decomposition through random-effect (GLS) regression. The regression model can be written as: $\tau_{sit} = \tau_t + u_s + e_{it}$, where subscript s is index for 4-digit industry, i for firm, t for year. τ_t is time fixed-effect, u_s is 4-digit industry random effect, e_{it} is firm-year idiosyncratic disturbance. Suppose $var(u_s) = s_u$, $var(e_{it}) = s_e$, and there is no correlation between u_s and e_{it}, then contribution of within-industry variance is measured by $((s_e)/(s_u + s_e))$. In STATA, we use the command: "bysort year: xtreg VAT-rate, i(4-digit-industry code)". And $((s_e)/(s_u + s_e)) = 1 - \rho$, where ρ can be obtained from running "xtreg".

Table 2: Sources of VAT Rate Dispersions

Dependent variable: VAT Rate τ_{it} (%) (Pooled Regression: 2000-2007)					
	(1)	(2)	(3)	(4)	(5)
Constant	10.66	10.65	10.78	10.76	16.92
	(0.04)***	(0.05)***	(0.04)***	(0.05)***	(0.18)***
SOE[1]	NO	NO	NO	NO	0.91
					(0.10)***
Size of firm[2]	NO	NO	NO	NO	-0.71
					(0.02)***
Ratio of Value-added[3]	NO	NO	NO	NO	-0.04
					(0.01)***
Export-Sales Rato[4]	NO	NO	NO	NO	-0.38
					(0.03)***
# of establishments[5]	NO	NO	NO	NO	0.01
					(0.01)
Year FE[6]	YES	YES	YES	YES	YES
Industry FE[7]	NO	YES	NO	NO	NO
County FE[8]	NO	NO	YES	NO	NO
Indu-County FE[9]	NO	NO	NO	YES	YES
R^2	0.00	0.06	0.11	0.44	0.55
Sample Size	899,265	899,265	899,265	899,265	374,649

Notes:
1. Dummy for SOE (State-owned-enterprise). SOE = 1 for State-owned-enterprise.
2. Size of a firm is measured by logarithm of its value-added.
3. Ratio of value-added, measured by value-added/(value-added + intermediate input), reflects the extent of a firm's vertical integration over its upstream business and products.
4. Export-sales ratio, measured by exports divided by a firm's sales.
5. Number of establishments within a firm.
6. Control for year fixed effects.
7. Control for (4-digit)industry fixed effects.
8. Control for county fixed effects.
9. Control for (4-digit)industry-county fixed effects.
10. Ordinary least squares. *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 3: Determinants of Input VAT Rate

Dependent variable: (Input VAT)/(Intermediate Goods Purchase) (%) (Pooled Regression: 2000-2007)					
	(1)	(2)	(3)	(4)	(5)
Constant	145.15	184.79	30.13	44.97	-828.32
	(42.65)***	(47.71)***	(30.72)	(42.16)	(152.51)***
SOE[1]	NO	NO	NO	NO	24.78
					(85.09)
Size of firm[2]	NO	NO	NO	NO	97.69
					(16.61)***
Ratio of Value-added[3]	NO	NO	NO	NO	5.6
					(6.84)
Export-Sales Rato[4]	YES	YES	YES	YES	-8.77
	NO	YES	NO	NO	(24.56)
# of establishments[5]	NO	NO	YES	NO	-6.73
	NO	NO	NO	YES	(10.75)
Year FE[6]	YES	YES	YES	YES	YES
Industry FE[7]	NO	YES	NO	NO	NO
County FE[8]	NO	NO	YES	NO	NO
Indu-County FE[9]	NO	NO	NO	YES	YES
R^2	0.00	0.00	0.53	0.63	0.79
Observation	898,343	898,343	898,343	898,343	373,987

Notes:
1. Dummy for SOE (State-owned-enterprise). SOE = 1 for State-owned-enterprise.
2. Size of a firm is measured by logarithm of its value-added.
3. Ratio of value-added, measured by value-added/(value-added + intermediate input), reflects the extent of a firm's vertical integration over its upstream business and products.
4. Export-sales ratio, measured by exports divided by a firm's sales.
5. Number of establishments within a firm.
6. Control for year fixed effects.
7. Control for (4-digit)industry fixed effects.
8. Control for county fixed effects.
9. Control for (4-digit)industry-county fixed effects.
10. Ordinary least squares. *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 4: Determinants of Output VAT Rate

Dependent variable: (Output VAT)/(Sales Turnover) (%) (Pooled Regression: 2000-2007)					
	(1)	(2)	(3)	(4)	(5)
Constant	10.42	10.79	10.57	10.86	12.39
	(0.06)***	(0.07)***	(0.07)***	(0.09)***	(0.18)***
SOE[1]	NO	NO	NO	NO	0.61
					(0.10)***
Size of firm[2]	NO	NO	NO	NO	-0.41
					(0.02)***
Ratio of Value-added[3]	NO	NO	NO	NO	5.39
					(0.01)***
Export-Sales Rato[4]	NO	NO	NO	NO	-0.48
					(0.03)***
# of establishments[5]	NO	NO	NO	NO	0.00
					(0.01)
Year FE[6]	YES	YES	YES	YES	YES
Industry FE[7]	NO	YES	NO	NO	NO
County FE[8]	NO	NO	YES	NO	NO
Indu-County FE[9]	NO	NO	NO	YES	YES
R^2	0.00	0.02	0.04	0.23	0.83
Observation	899,250	899,250	899,250	899,250	374,649

Notes:
1. Dummy for SOE (State-owned-enterprise). SOE = 1 for State-owned-enterprise.
2. Size of a firm is measured by logarithm of its value-added.
3. Ratio of value-added, measured by value-added/(value-added + intermediate input), reflects the extent of a firm's vertical integration over its upstream business and products.
4. Export-sales ratio, measured by exports divided by a firm's sales.
5. Number of establishments within a firm.
6. Control for year fixed effects.
7. Control for (4-digit)industry fixed effects.
8. Control for county fixed effects.
9. Control for (4-digit)industry-county fixed effects.
10. Ordinary least squares. *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 5: VAT Rate and Reform - Summary Statistics

	All Provinces		Northeast Provinces[1,2]		Other Provinces[3]	
year	mean (%)	S.D. (%)	mean (%)	S.D. (%)	mean (%)	S.D. (%)
2000	10.7	11.8	12.0	14.0	10.6	11.6
2000	11.0	11.9	11.0	12.6	11.0	11.9
2002	10.5	11.0	11.1	12.5	10.5	10.9
2003	10.5	10.9	10.7	13.8	10.4	10.7
2004	10.3	11.5	10.8	14.4	10.3	11.2
2005	10.2	11.3	9.0	14.1	10.3	11.1
2006	10.0	11.1	8.0	11.8	10.2	11.0
2007	10.0	10.5	8.0	10.1	10.1	10.5

Notes:

1. Northeast provinces include Jilin, Liaoning, and Heilongjiang.

2. Consumption-type VAT was only applicable in six selected industries, including agricultural product processing, equipment manufacturing, petrochemical, metallurgy, ship building, automobile manufacturing, military products and high-tech products.

3. In July of 2007, the reform was extended to another twenty-six cities in six central-region provinces. In this column these cities are not dropped from "Other Provinces" due to very short sample period (only 6 months).

Table 6: VAT Rate (%) and Reform - Regression [1]

	τ_{it}	$Std(\tau_{it})$
	(1)	(2)
$R_i^{NE} I_i^s Y_t^{After}$	-0.76	-1.25
	(0.60)	(0.58)**
$R_i^{NE} Y_t^{After}$	-2.21	-0.32
	(0.41)***	(0.19)*
$I_i^s Y_t^{After}$	0.49	0.29
	(0.29)*	(0.19)
$R_i^{NE} I_i^s$	-0.05	1.85
	(1.21)	(0.82)**
$R_i^{NE,2}$	0.62	0.61
	(0.46)	(0.23)***
$I_i^{s,3}$	-1.36	0.44
	(0.54)**	(0.19)**
$Y_t^{After,4}$	-0.39	-0.81
	(0.14)***	(0.07)***
Constant	10.85	7.3
	(0.20)***	(0.08)***
R^2	0.00	0.01
Observation	899,265	69,148

Notes:

1. The regression is based on triple difference (Equation (4.3)) - between regions, between industries, over time periods - driven by the consumption-type VAT reform in Northeast provinces in 2004. τ_{it} is the effective VAT rate at firm level. $Std(\tau_{it})$ is the standard-deviation of τ_{it} within (2-digit)industry-province cells.

2. Regional dummy R_i^{NE} for Northeast provinces, including Jilin, Liaoning, and Heilongjiang.

3. Industrial dummy I_i^s for six industries under consumption-type VAT reform, including agricultural product processing, equipment manufacturing, petrochemical, metallurgy, ship building, automobile manufacturing, military products and high-tech products.

4. Time dummy Y_t^{After} for years post reform (after January of 2005). In July of 2007, the reform was extended to another twenty-six cities in six central-region provinces, which are not taken into account due to very short sample period (only 6 months).

5. Ordinary least squares. *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level. Robust standard errors are clustered at the (2-digit)industry-province level (1,177 groups).

Table 7: VAT Rate Dispersions and Export Refund

	Ex-post Refund[1]		Ex-ante Refund[2]	
year	mean (%)	S.D. (%)	mean (%)	S.D. (%)
2000	10.7	11.8	10.8	11.8
2001	11.0	11.9	11.1	12.0
2002	10.5	11.0	10.7	11.0
2003	10.5	10.9	10.6	10.9
2004	10.3	11.5	10.4	11.4
2005	10.2	11.3	10.3	11.3
2006	10.0	11.1	10.2	11.1
2007	10.0	10.5	10.1	10.5

Notes:
1. This column lists the mean and standard deviation of the effective VAT rates in the Annual Survey of Industrial Production.
2. This column lists the mean and standard deviation of the counter-factual VAT rates with export refund not being conducted. To get the counter-factual VAT rates, I run the following regression: $VAT\_rate = \alpha + \beta * (Export/VA) + \gamma * X + \epsilon$, where $Export$ is firm's export value, VA is firm's value-added, and X are a vector of firm's characteristics, including county, 4-digit industry, and ownership type. The counter-factual VAT rate is equal to "Effective VAT rate $-\hat{\beta} * (Export/VA)$" where $\hat{\beta}$ is estimator for β.

Table 8: Effect of Entry and Exit of firms on VAT Rate

Dependent Variable[1]	
$Ave\_\tau_{st}^{G2} - Ave\_\tau_{st}^{G1}$ (%)[2]	$Std\_\tau_{st}^{G2} - Std\_\tau_{st}^{G1}$ (%)[3]
-0.13*[4]	1.02***[4]
(-1.68)	(9.88)

Notes:
1. A dependent variable is regressed on a constant, which indicates the gap between the Group 1 and the Group 2.
2. $Ave\_\tau_{st}^{Gi}$ is the average of VAT rate over firms of group i in industry s in year t.
3. $Std\_\tau_{st}^{Gi}$ is the standard deviation of VAT rate over firms of group i in industry s in year t.
4. Estimated constant. *** (**, *) indicates statistical significance at the 99%(95%, 90%) confidence level.

Table 9: Serial Correlation of VAT Rate over Years

	Sample 1[1]		Sample 2[2]	
	2001-2002	2002-2003	2005-2006	2006-2007
Firm Level	0.53	0.62	0.60	0.60
County Level	0.78	0.80	0.81	0.83
City Level	0.89	0.89	0.90	0.92
Province Level	0.96	0.97	0.96	0.97
4-digit Industry Level	0.77	0.49	0.90	0.75
2-digit Industry	0.97	0.95	0.99	0.99

Notes:

1. Because of the change in firm ID, we can not link the firms before 2004 to firms thereafter. Year 2004 is not included due to inconsistence of firm ID. Sample 1 only includes the firms staying in the survey for 2001, 2002, and 2003 period. The sample size is 50,236. Year 2000 is not included so that Sample 1 and 2 are both comparably three years.

2. Sample 2 only includes the firms staying in the firm survey for 2005, 2006, 2007 period. The sample size is 66,345.

Table 10: Correlation between VAT Rate and Firm Size [1]

Value-added	Capital	Labor	ln(Value_added)	ln(Capital)	ln(Labor)
0.0079	0.0221	0.0184	-0.0225	0.0962	0.029

Notes:

1. Size of a firm is measured respectively by its value-added, capital stock, number of employee, and their logarithm, including year 2000, 2001, 2002, 2003, 2005, 2006, 2007.

Table 11: Summary Statistics–Firm Survey (2000 - 2007) [1]

Variable	Mean	Std.Dev	Min	p1[2]	p50[2]	p99[2]	Max
Capital	108031.6	1450939	1	982	14285	1523384	5.64E+08
Labour	216.1752	949.4027	1	10	82	2183	158288
Value-added	38123.7	462637	4	708	8199	443205	1.63E+08
Wage Bill	14485.27	145675.7	3.6	212.4	3247.2	171399.6	3.02E+07
VAT Rate[3]	0.1029848	0.111686	-4.364841	0	0.081578	0.4647059	10.96596

Notes:
1. Sample size = 899,265. There are 66,222 firms in year 2000, 74,511 in year 2001, 83,285 in year 2002, 97,993 in year 2003, 131,306 in year 2004, 129,398 in year 2005, 147,696 in year 2006, 168,876 in year 2007.
2. p1, p50, p99 respectively are 1-*th*, 50-*th*(median), 99-*th* percentile.
3. The (Effective) VAT Rates are calculated by dividing the payable value-added tax by the value-added in each firm.

Table 12: Summary Statistics - Population Census (2005)

Population Census[1]	Mean	Std. Dev.	Min.	Max.
Wage	11830.86	12213.15	12	1199988
Education	9.471699	2.746082	0	19
Age	34.39995	10.33802	18	70

Notes:
1. The data is from 1% sampling Population Census Conducted by the National Bureau of Statistics of China in Year 2005. Sample size =208,153.

Table 13: Dispersions in VAT and TFP/GDP Loss

Year	2000	2001	2002	2003	2004	2005	2006	2007
(1):TFP Loss (%)[1]	58.7	54.5	70.3	72.8	62.6	67.3	67.1	78.5
(2):TFP Loss Net of VAT (%)[2]	56.7	52.6	68.2	71.0	59.9	65.4	65.0	75.1
(3):Contribution of VAT (%)[3]	5.0	4.3	7.4	6.5	7.4	5.7	6.6	15.9
(4):VAT Rate (%)[4]	10.7	11.0	10.5	10.5	10.3	10.2	10	10.0

Notes:
1. This row is the TFP loss caused by all distortions based on HK(2009), with the elasticity of substitution σ equal to 3.
2. This row is the TFP loss caused by all the distortions net of VAT.
3. This row is the contribution of VAT dispersion to the TFP loss, which is equal to | (1- TFP Loss Net of VAT)/(1-TFP Loss) -1 | $\times$100%.
4. This row is the total payable VAT divided by the total value added in the whole manufacturing sector.

Table 14: Correlation between Distortions and Productivity

	$ln(1-\tau_{si}^{VAT})$	$ln(1+\tau'_{k_{si}})$	$ln(1+\tau'_{H_{si}})$	$TFPR_{si}^{KH,1}$	lnA_{si}
$ln(1-\tau_{si}^{VAT})$	1.0000				
$ln(1+\tau'_{k_{si}})$	-0.0761	1.0000			
$ln(1+\tau'_{H_{si}})$	0.1943	0.1503	1.0000		
$TFPR_{si}^{KH,1}$	-0.0009	0.9141	0.4935	1.0000	
lnA_{si}	-0.0255	0.3711	0.6251	0.5079	1.0000
Std. Dev. [2]	0.1451	1.6060	1.0398	1.1036	1.3000

Notes:
1. $TFPR_{si}^{KH} = \alpha_s ln(1-\tau'_{k_{si}}) + (1-\alpha_s) ln(1-\tau'_{H_{si}})$.
2. Standard deviation of variables.

Table 15: TFP Loss with VAT Tax Refund Excluded [1]

Year	2000	2001	2002	2003	2004	2005	2006	2007
(1):Contribution of VAT (%)[2]	3.0	4.2	9.5	6.6	7.3	6.0	7.8	16.0
(2):Benchmark contribution (%)[3]	5.0	4.3	7.4	6.5	7.4	5.7	6.6	15.9
(3):VAT Rate net of Refund (%)[4]	10.8	11.0	10.7	10.6	10.3	10.3	10.2	10.1
(4):Benchmark VAT Rate (%)[5]	10.7	11.0	10.5	10.5	10.3	10.2	10.0	10.0

Notes:
1. The VAT rate used in this table is the counter-factual VAT rate with export tax refund being excluded.
2. This row is the contribution of VAT dispersion to the TFP loss.
3. This row is the contribution of VAT dispersion to the TFP loss in Table 13 (benchmark).
4. This row is the total payable VAT divided by the total value added in the whole manufacturing sector.
5. This row is the total payable VAT divided by the total value added in the whole manufacturing sector in Table 13 (benchmark).

Table 16: Time Average of VAT Rate (%, 2000-2003)

Year	2000[5]	2001[5]	2002[5]	2003[5]	Year Average[6]
(1):Sample 1[1]	1.8	1.9	5.5	1.5	1.5
(1):Sample 2[2]	2.1	2.0	7.4	6.3	2.2
(3):Sample 3[3]	5.6	3.5	7.9	6.8	3.4
(4):Sample 4[4]	5.0	4.3	7.3	6.5	4.0

Notes:
1. Sample 1 only includes firms staying over year 2000, 2001, 2002, and 2003.
2. Sample 2 includes firms staying for at least three years in year 2000, 2001, 2002, and 2003.
3. Sample 3 includes firms staying for at least two years in year 2000, 2001, 2002, and 2003.
4. Sample 4 is the full sample.
5. The results in this column are calculated with corresponding year's effective VAT rate for firms in the corresponding sample.
6. The results in this column are calculated with the year-average effective VAT rate for firms in the corresponding sample. Firms staying in the sample for three years take the 3-year average of VAT rate. And firms staying for two years take 2-year average. Firms staying for only one year takes its effective VAT rate in that year.

Table 17: Time Average of VAT Rate (%, 2005-2007)

Year	2005[4]	2006[4]	2007[4]	Year Average[5]
(1):Sample 5[1]	3.7	4.0	12.6	4.0
(2):Sample 6[2]	3.9	5.3	13.3	4.2
(3):Sample 7[3]	5.7	6.6	15.9	9.2

Notes:
1. Sample 5 only includes firms staying over year 2005, 2006, and 2007.
2. Sample 6 includes firms staying for at least two years in year 2005, 2006, and 2007.
3. Sample 7 is the full sample.
4. The results in this column are calculated with corresponding year's effective VAT rate for firms in the corresponding sample.
5. The results in this column are calculated with the year-average effective VAT rate for firms in the corresponding sample. Firms staying in the sample for three years take the 3-year average of VAT rate. And firms staying for two years take 2-year average. Firms staying for only one year takes its effective VAT rate in that year.

Table 18: Data Quality [1]

Year	2000	2001	2002	2003	2004	2005	2006	2007
(1):Contribution of VAT (%)[2]	6.8	4.3	3.5	2.7	7.0	4.6	4.5	21.3
(2):Benchmark contribution (%)[3]	5.0	4.3	7.4	6.5	7.4	5.7	6.6	15.9
(3): VAT Rate (%)[4]	10.8	11.0	10.5	10.2	9.9	10.2	10.2	10.0
(4):Benchmark VAT Rate (%)[5]	10.7	11.0	10.5	10.5	10.3	10.2	10.0	10.0

Notes:

1. The sample here only includes the firms in those industries where VAT Tax Burden Ratio lie within the warning bounds issued in 2007 by the National Bureau of Taxation.

2. This row is the contribution of VAT dispersion to the TFP loss.

3. This row is the contribution of VAT dispersion to the TFP loss in Table 13 (benchmark).

4. This row is the total payable VAT divided by the total value added in the whole manufacturing sector.

5. This row is the total payable VAT divided by the total value added in the whole manufacturing sector in Table 13 (benchmark).

Table 19: TFP Loss with Outliers Excluded [1]

Year	2000	2001	2002	2003	2004	2005	2006	2007
(1):Contribution of VAT (%)[2]	5.0	4.3	7.4	6.6	7.4	5.8	4.4	6.5
(2):Benchmark contribution (%)[3]	5.0	4.3	7.4	6.5	7.4	5.7	6.6	15.9
(3):VAT Rate (Average, %)[4]	10.6	10.9	10.5	10.4	10.3	10.1	10.0	9.9
(4):VAT Rate (Min, %)[5]	-16.8	-23	-25.4	-30	-44.5	-30.7	-31	-28.6
(5):VAT Rate (Max, %)[6]	100	107.2	100	100	105.8	100.1	98.2	98.9
(6):Benchmark VAT Rate (%)[7]	10.7	11.0	10.5	10.5	10.3	10.2	10.0	10.0

Notes:

1. The sample here excludes the firms with the effective VAT rates lower than 1-*th* percentile or greater than 99-*th* percentile. 2. This row is the contribution of VAT dispersion to the TFP loss.

3. This row is the contribution of VAT dispersion to the TFP loss in Table 13 (benchmark).

4. This row is the total payable VAT divided by the total value added in the whole manufacturing sector.

5. This row is the firms' minimum effective VAT rate in each year.

6. This row is the firms' maximum effective VAT rate in each year.

7. This row is the total payable VAT divided by the total value added in the whole manufacturing sector in Table 13 (benchmark).

Table 20: TFP Loss with $\sigma = 5$ [1]

Year	2000	2001	2002	2003	2004	2005	2006	2007
(1):Contribution of VAT (%)[2]	14.8	19.0	26.7	23.5	40.0	34.9	36.7	58.3
(2):Benchmark contribution (%)[3]	5.0	4.3	7.4	6.5	7.4	5.7	6.6	15.9

Notes:

1. The results in this table are based on calculation with the elasticity of substitution σ equal to 5.

2. This row is the contribution of VAT dispersion to the TFP loss.

3. This row is the contribution of VAT dispersion to the TFP loss in Table 13 (benchmark).

Table 21: TFP Loss with VAT Reform Provinces and Industries Excluded [1]

Year	2000	2001	2002	2003	2004	2005	2006	2007
(1):Contribution of VAT (%)[2]	5.8	4.8	7.2	8.5	6.7	5.7	5.2	8.8
(2):Benchmark contribution (%)[3]	5.0	4.3	7.4	6.5	7.4	5.7	6.6	15.9
(3):VAT Rate (%)[4]	11.0	11.3	10.8	10.8	10.5	10.5	10.5	10.4
(4):Benchmark VAT Rate (%)[5]	10.7	11.0	10.5	10.5	10.3	10.2	10.0	10.0

Notes:

1. The sample here excludes all firms in eight industries AND three north-east provinces under 2004 VAT reform.

2. This row is the contribution of VAT dispersion to the TFP loss.

3. This row is the contribution of VAT dispersion to the TFP loss in Table 13 (benchmark).

4. This row is the total payable VAT divided by the total value added in the whole manufacturing sector.

5. This row is the total payable VAT divided by the total value added in the whole manufacturing sector in Table 13 (benchmark).

Table 22: TFP Loss Caused by Distortions in K and H [1]

Distortions	K	H	(K, H)
TFP Loss (%, Our method[2])	11	13	19
TFP Loss (%, HK(09) method[3])	11	9	14

Notes:

1. To make comparison, the TFP loss is calculated based on the industry-province-ownership "firm-cell", rather than the real firms in firm survey. There are thirty-nine 2-digit industries, thirty-one provinces, and two types of firm ownership (SOE, and Non-SOE) and 1,860 "firm-cell" in our sample.

2. Our method uses hedonic wage as the proxy for human capital for each "firm-cell".

3. HK(09) method uses wage bill as the proxy for human capital for each "firm-cell" (Hsieh and Klenow, 2009).