

Do Government Accountability and Responsiveness Enhance Support for Property Taxes? Experimental Evidence from China*

Xiaobo Lü¹, Lily Tsai², Ping Zhang³ and Minh Trinh⁴

¹University of Texas at Austin

²Massachusetts Institute of Technology

³Fudan University

⁴Purdue University

Version: 4/8/2025

Abstract

The adoption of property taxes is often met with significant challenges in developing countries. Enhancing administrative capacity for assessment and compliance is undoubtedly important, but addressing the political impasse caused by potential mass resentment poses a more formidable challenge. We explore whether institutions for responsiveness and accountability enhance tax morale in societies with weak democratic institutions. We investigate two dimensions of good governance that the government can uphold: government responsiveness to citizen demands in policymaking; and holding corrupt officials accountable. To assess their impact on property tax preferences and compliance behaviors in urban China, we conducted a conjoint experiment and a lab-in-the-field experiment. The former reveals that both responsiveness and accountability have positive effects on preferences for property tax policies; however, the results of the latter provide strong evidence that top-down institutions of accountability are more influential in shaping tax compliance behavior.

Keywords: Property Tax; Governance; Experiment; China

*This research received financial support from the Samuel Tak Lee Real Estate Entrepreneurship Laboratory at the Massachusetts Institute of Technology. We thank Jennifer Pan, Yiqing Xu, Justin Grimmer, Vicky Fouka, and Joseph Wang as well as participants in seminars at New York University, the Hoover Institution and the 2016 and 2018 APSA conference for helpful comments at different stages of the study. My Seppo, Nina McMurry, Jing Cao, Xiaohui Jiang, Zheng Liu, Peixuan Huang, and Zhuoqi Yu as well as numerous Fudan students provided excellent research assistance. All errors remain our own.

[†]Corresponding author. Please email xiaobolu@austin.utexas.edu with questions and comments.

1 Introduction

Property taxes, particularly those levied annually on property owners, have long served as a vital source of state revenue compared to other forms of taxation because of its appealing attributes: fairness, efficiency, and stability.¹ Despite these appealing qualities, property taxes has yet to emerge as a major source of revenue in the developing world. By the 2010s property tax revenue accounted for a mere 0.51% of the GDP in developing and transitional countries in stark contrast to the 1.82% observed in OECD countries.² Conventional wisdom often attributes the obstacles to adopting property taxes to daunting administrative and enforcement costs;³ however, the visibility inherent in property tax also makes it a salient issue, forcing the state to approach it cautiously for fear of potential backlash. Stated differently, many countries are likely to be *reluctant* rather than *incapable* of adopting property tax.

The political sensitivity surrounding property tax can best be exemplified by the Chinese government's failed attempts to tax private homeownership in urban areas over the past two decades. By many measures, China should have encountered fewer hindrances to adopting property tax than other developing countries because of its robust administrative capacity for taxation⁴ and its dominant ruling party (the Chinese Communist Party, or CCP hereafter).⁵ As part of its market-oriented reforms launched in the 1980s, China implemented a series of housing policy reforms in urban areas, leading to a booming real estate market and significant private ownership of urban properties.⁶ Following the CCP's declaration of its intention to institute a property tax on urban private properties in 2003, the Chinese government embarked on a series of initiatives laying the groundwork for property tax collection.⁷

¹See Collier et al. (2017) for a discussion of the benefits of property tax collection.

²Calculated by the authors based on IMF Global Revenue Statistics Database (<https://data.imf.org/?sk=77413f1d-1525-450a-a23a-47aee40fe78>)

³Norregaard (2013) highlights the potential and challenges in implementing property tax.

⁴Cui (2022) offers an excellent study tracing the development of China's taxation administrative capacity since the 2000s.

⁵For the expansion of state and party power in the Chinese economy in the last two decades, see Lardy (2019) and Pearson, Rithmire and Tsai (2023).

⁶See Man (2011) for an overview of China's housing reform.

⁷See Zhang (2020) for a discussion of the evolution of property tax in China and the origins of the pilot programs in Shanghai and Chongqing.

For instance, Shanghai and Chongqing pioneered pilot programs for annual property tax collection targeting wealthier individuals in 2011 (Hou, 2018), followed by regulations aimed at establishing a national homeownership registration database in 2016.⁸ Nonetheless, the expansion of property tax proves to be challenging: Chongqing and Shanghai remain the sole cities collecting property tax to this date. Even President Xi Jinping, who is arguably the most dominant Chinese leader since Deng Xiaoping, attempted but failed to expand the property tax in 2021 for fear of potential political and economic impacts.⁹

The challenge of introducing property tax in China highlights its complex role, affecting both macroeconomic growth and state–society relations. Private homeownership has been a major source of household wealth in urban China, accounting for 63.9% of household wealth in 2022 in contrast to 28.5% in the United States;¹⁰ therefore, Chinese policymakers are acutely concerned about how citizens might respond to a broad-based property tax. Given China’s long history of social unrest over taxation—most recently, some violent rural protests, resulting in the abolition of rural taxation in 2005—convincing citizens to support new demands for tax payment is a core political challenge.¹¹

How can the government mitigate political resistance to property tax and achieve compliance from potential taxpayers? Surprisingly, extant studies have offered few answers to these questions in the case of China.¹² Prior research on the prospective implementation of property tax in China has primarily centered on its potential ramifications in the macroecon-

⁸The Ministry of National Resources announced detailed regulations on property registrations in 2016.

⁹Xi made a bold attempt to introduce property tax collection in major cities as part of his Common Prosperity agenda in autumn 2021, however, this plan met resistance within the party and was subsequently canceled because of the difficult macroeconomic conditions exacerbated by the COVID-19 pandemic. Introducing the property tax during economic downturns exacerbate the worsening economy, directly undermining the CCP’s popular support based on performance legitimacy. See “In Tackling China’s Real-Estate Bubble, Xi Jinping Faces Resistance to Property-Tax Plan,” *Wall Street Journal* (Oct. 19, 2021).

¹⁰See “2022 Trends in China Household Wealth” (<https://chfs.swufe.edu.cn/info/1031/1736.htm>) produced by the Survey and Research Center for China Household Finance, Southwest University of Finance and Economics. U.S. household wealth composition is based on “The Wealth of Households: 2021” produced by the United States Census Bureau (<https://www.census.gov/library/publications/2023/demo/p70br-183.html>).

¹¹See Bernstein and Lü (2003) and Bianco (2001) on rural taxation and protests in China. For the abolition of rural taxation in 2005, see Göbel (2010).

¹²See Hou, Ren and Zhang (2014) for a review of the history and prospects of property tax in China.

omy,¹³ the real estate sector,¹⁴ revenue contributions to local government,¹⁵ and homeowners' capacity to bear the tax burden and income redistribution.¹⁶ Although a handful of scholars have explored Chinese citizens' attitudes toward property tax, their research has largely focused on the economic aspects of property taxes (i.e., tax rates, exemption criteria, the use of property tax revenue, and reminders for overdue payment).¹⁷

We contend that citizen support for a tax policy does not rely solely on economic calculations but also hinges on tax morale, by which nonpecuniary factors enhance tax compliance. Increasingly, scholars have recognized the importance of tax morale to facilitate tax compliance. An important facet of tax morale is rooted in the concept of reciprocity, where voluntary tax compliance is viewed as an exchange for political or economic benefits, such as political representation and the expansion of government services. Demanding political representation, however, is a costly endeavor—economic elites, not ordinary citizens, are often the ones at the forefront in demanding political representation in exchange for their taxation.¹⁸ Meanwhile, the notion that the delivery of government service can enhance tax compliance has garnered mixed empirical support in extant experimental studies.¹⁹

In this paper, we shift our focus to another dimension of reciprocity that the state could offer in return for support of property tax adoption—upholding government responsiveness and accountability—which in turn enhances state legitimacy and policy compliance.²⁰ Citizens are more likely to comply voluntarily with state demands when they are legitimate²¹

¹³Yang and Liu (2023) and Zhou and Pan (2021).

¹⁴Xu, Yin and Wang (2023) and Zhang, Hou and Li (2020).

¹⁵Li (2014), Liu, Zhang and Chen (2015), and Zhang, Yao and Feng (2021).

¹⁶Cao and Hu (2016), Zhang and Hou (2016), and Zhan and Li (2015).

¹⁷See, for example, Antinyan et al. (2021), Liu, Chen and Wei (2020), and Liu, Gao and Tao (2019). The rationale behind this strategy is motivated by the classic economic model of tax compliance, where the tax rate, the detection probability, and the penalty for tax evasion are the primary parameters shaping tax compliance (Allingham and Sandmo, 1972).

¹⁸A close reading of canonical scholarship on fiscal bargaining based on early modern Europe reveals that economic elites were at the forefront of fiscal bargaining (e.g., Bates and Lien (1985), North and Weingast (1989), and Levi (1989)). Kao, Lü and Queralt (2024) offer direct evidence based on a conjoint experiment for hypothetical tax reforms in separate samples of business elites and ordinary citizens in China.

¹⁹See Luttmer and Singhal (2014) for a summary of the literature.

²⁰Tsai (2021), Tyler (2006)

²¹See Levi (1989), Levi and Sacks (2009), Murphy, Tyler and Curtis (2009).

and lead to good governance.²² One key aspect of good governance is responsiveness to grassroots citizen demands in policymaking. This bottom-up approach has deep roots in the scholarship of democratic theory and the social contract: Citizens are more likely to comply with taxation, or government policies in general, when the state incorporates and responds to their preferences, increasing the likelihood of their perception the process of policymaking as fair or procedurally just. Ultimately, such policies are better for public welfare.²³ A second key aspect of good governance is accountability for corruption and incompetence. Citizens may be more likely to comply with taxation and other government demands when the state punishes politicians with poor performance; hence citizens are more likely to perceive the government as committed to public values and retributive justice, and public funds may be utilized more effectively too.²⁴

To test our argument, we employed a survey experiment and a lab-in-the-field experiment to evaluate both citizens' preferences and behaviors regarding property taxes in China. We first carried out a nation-wide conjoint experiment to evaluate the relative strength of classic economic factors vis-à-vis signaling government responsiveness and accountability in the design of property tax policies. The results of our conjoint experiment reveal that Chinese respondents value institutional design for both top-down sanctioning and bottom-up citizen input. Importantly, the marginal effects of these two attributes resemble the classic economic factors signaling the benefits and costs of tax compliance, suggesting institutions for accountability and responsiveness are as effective as the parameters in the standard economic model of tax compliance.

Motivated by the revealed *preference* in the conjoint experiment, we conducted a lab-in-the-field experiment in Shanghai to evaluate tax compliance *behaviors* when subjects have the opportunity to observe government responsiveness and accountability in real time. We employed an iterated public goods game, in which participants made decisions concerning

²²See, for example, Rothstein (2009).

²³See Levi and Stoker (2000) for a review.

²⁴Tsai (2021).

compliance with property tax payments and funding local public goods and services. Respondents played several rounds of this game, and in later rounds they received a treatment informing them whether the relevant authority figure had or had not been responsive to their preference for the use of property tax revenue (i.e., the responsiveness treatment) or punished corrupt officials (i.e., the accountability treatment). Our experiment revealed an important finding.

We found strong evidence that the accountability treatment consistently enhanced tax compliance but only very weak evidence for the responsiveness treatment. In the round when the subjects were first exposed to the experimental treatment, we observed a sharp increase in tax compliance—18.5 percentage points for the accountability treatment. The positive treatment effects persisted in subsequent treatment rounds and ranged between 11.6 percentage points and 17.8 percentage points, depending on model specification. By contrast, the responsiveness treatment had little impact on our subjects' tax compliance decision. Our findings remained robust across various additional analyses.

Our paper sheds light on two important strands of literature. First, this insight holds essential implications for societies where democratic structures are either absent or ineffective. Conventional wisdom has emphasized improvement of bureaucratic capacity or incentives to enhance state taxation capacity,²⁵ yet focusing solely on capacity building can be challenging and inadequate in this context, where resources and political will are often in short supply in developing countries. More importantly, the primary challenges facing many developing nations do not necessarily result from a lack of administrative capacity but instead from the potential political consequences that could undermine state legitimacy and popular support. Our findings illustrate that promoting government legitimacy by strengthening top-down bureaucratic sanctioning is a crucial strategy boosting tax morale that complements building administrative capacity.

²⁵The rationale stemming from the economic theories of compliance suggests that taxpayers "play the audit lottery," which maximizes their expected after-tax return—possibly after adjustment for the desired level of risk (Trivedi, Shehata, and Mestelman, 2005:2). See Slemrod (2019) for a review of the field.

This study joins recent studies on the tax compliance issue in the developing world by employing experimental methods. Our paper differs in two important aspects. First, we evaluate whether citizens' perception of good governance enhances their tax compliance behavior²⁶ but not how taxation incentivizes citizens' political engagement and demand for government accountability.²⁷ Second, we study both top-down and bottom-up mechanisms of good governance. Contrary to existing studies focusing on bottom-up representation and accountability, our findings suggest that responding to citizen demands for government services does not significantly boost tax morale in China. Instead, signaling top-down accountability (e.g., retributive justice), such as through initiatives like Xi's anticorruption campaign, may alleviate resistance to new taxes and bolster compliance.

In the next section, we lay out our theoretical framework for examining the impact of grassroots participation and bureaucratic sanctioning on citizen willingness to comply with taxation. We then present our empirical strategy and discuss the experimental results. We conclude by discussing the implications of our findings and suggesting directions for future research. Specifically, we underscore the broader context and additional conditions that influence citizens' support for property taxes, drawing on our findings.

2 Government Responsiveness, Accountability, and Policy Compliance

Citizens' voluntary policy compliance often stems from their perception of whether or not the state upholds responsiveness to citizens and accountability for good governance.²⁸ Two

²⁶Previous studies have shown that lack of access to government services undermines tax morale. See Ali, Fjeldstad and Sjurson (2014), Bodea and LeBas (2014), and Castañeda, Doyle and Schwartz (2020). Meanwhile, Flores-Macías (2018) found some evidence that citizen oversight improved support for taxation. Finally, Timmons and Garfias (2015) shows that corruption undermined municipal property tax revenue in Brazil.

²⁷See, for example, de la Cuesta et al. (2022), Martin (2023), Paler (2013), and Weigel (2020). For the exception of indirect taxation, see de la Cuesta et al. (2023).

²⁸See, for example, Rothstein (2009), Murphy, Tyler and Curtis (2009), and Tyler (2006).

distinct facets of good governance—*responsiveness* and *accountability*—could have profound impact on citizen behaviors in policy compliance and are rooted in the canonical theories of the social contract and social justice. To begin, the state could maintain *responsiveness* by addressing citizen demands through grassroots political participation. When citizens believe that authorities employ equitable procedures in formulating policies and responding to citizen demands—that is, they uphold procedural justice—they are more inclined to comply with those decisions.²⁹ Applying this line of logic to tax compliance, the prevailing view suggests that grassroots participation in policymaking enhances tax compliance because citizens are more likely to pay taxes when the government uses a legitimate process for making decisions about how to allocate public resources and responds to their preferences about the way their taxes are spent. The basic premise is that citizens believe that the government should deliver promised and fair returns for their taxes.

In the developing world, however, using institutions of procedural justice to encourage policy compliance might prove futile. This is often the result of uncertainty and ineffectiveness in the decision-making procedures of the government, a function of the weak formal institutions that are often common in the developing world. Canonical theories are built on the studies of consolidated democracies, where the rule of law is strong and goes hand in hand with political rights and participatory institutions. In these political systems, democratic elections together with the rule of law ensure that citizens can select officials who care about the public interest and punish officials who fail to promote it. In nondemocratic and hybrid regimes, however, citizens have no guarantee that government authorities will respond to their input.

Uncertainty surrounding the effectiveness of citizen action and political fear are critical obstacles to grassroots participation in developing contexts for several reasons. First, formal procedures for government decision making are often unclear or unknown to ordinary people. Second, officials often have an interest in making the workings of government as

²⁹See Dahl (1956), Cohen-Charash and Spector (2001), and De Cremer and Tyler (2005).

“remote and incomprehensible” as possible.³⁰ Finally, citizens in these contexts may also refrain from providing input—even through formally sanctioned channels—out of fear or because they perceive such efforts to be useless. “When citizens perceive the bureaucracy as insensitive,...these perceptions may lead to diminished belief in the value of citizens’ involvement and political efficacy, negatively affecting citizens’ willingness to participate in politics” (Vigoda-Gadot, 2007, 290).

To encourage perceptions of legitimate authority and good governance, the state can highlight its efforts to uphold government accountability by sanctioning officials for poor performance (Fearon, 1999, 55). Top-down institutions for accountability, such as anticorruption measures, may increase citizen willingness to pay taxes because citizens believe that the taxes they pay will contribute to public welfare rather than being wasted or misappropriated by corrupt officials (Field and Frey, 2007). The top-down bureaucratic sanctioning of incompetent officials can serve as a signal to taxpayers, indicating that the noncompliant among them will face consequences, thus deterring free-riding behavior and incentivizing tax compliance (Levi and Sacks, 2009). Higher levels of corruption can lead citizens to conclude that their taxes are being squandered, and they may decide they have no reason to uphold their part of the bargain if the government is not upholding its (Braithwaite, 1998).

Accountability and the punishment of corruption may also increase tax morale by signaling the state’s commitment to public integrity and retributive justice (Tsai, 2021; Tsai, Trinh and Liu, 2022).³¹ When authorities punish corruption, they show themselves to be moral actors and leaders (Blau, 1964). Citizens believe that these leaders uphold moral standards and care about the good of the community. In consolidated democracies, retributive justice is embedded into the rule of law. Citizens in liberal democracies are confident (relatively speaking) that the law and the judiciary system exist to punish misbehaving officials. For nondemocratic and hybrid regimes that seek to be seen by citizens as at least somewhat fair and just, institutions such as anticorruption campaigns and anticrime initiatives can create

³⁰Khanna and Johnston (2007).

³¹For studies beyond China, see Carlsmith and Darley (2008) and von Hirsch et al. (2003).

citizen perceptions of retributive justice, which may account in large part for the popularity of political leaders like Xi Jinping in China, Duterte in the Philippines, and Putin in Russia.

3 Research Design

China is an ideal testing ground to evaluate government responsiveness and accountability in regimes with weak institutions for two reasons. First, citizens in China, similar to many others in the developing world, have limited options to influence policymaking, even though scholars have found that local governments are responsive to citizen input in some instances.³² Second, government accountability through top-down sanctioning may lack credibility in nondemocracies; however, since Xi Jinping launched the anticorruption campaign in 2012, the credibility of such accountability measures has significantly improved in China. Hence, both top-down accountability and bottom-up grassroots participation, while far from perfect, convey a degree of credibility within the Chinese context.

Our research design consisted of two stages.³³ In the first stage we conducted a nationwide conjoint experiment across major Chinese cities to assess citizen *preferences* for property tax policies. The experiment presented respondents with pairs of randomized tax policy designs that varied across several dimensions, thus analyzing respondents' choice in each pair yielded aggregate preferences given the institutional design of accountability and responsiveness in tax policy.

In a political environment with weak institutions, the actual policy implementation may not always reflect the institutional design. In the second stage we carried out a lab-in-the-field experiment to investigate whether *experiencing* government responsiveness and accountability enhances citizen compliance *behavior* in the city of Shanghai. This experiment simulated the property tax environment with an adapted, iterated public goods game, in which each player decided whether to contribute their property tax payment to a pool of common re-

³²See Chen, Pan and Xu (2016) and Distelhorst and Hou (2017).

³³Both survey and lab-in-the-field experiments were reviewed and granted an exemption by the Human Subjects Review Committee at the authors' respective institutions.

sources from which everyone would receive a payout or to free-ride on other players' contributions. Halfway through the iterated game, we introduced experimental manipulations to simulate whether the authority in charge of disbursing payouts upholds or disavows accountability or responsiveness. Specifically, we compared respondents' decisions in tax compliance before and after observing manipulated government officials' behaviors in accountability or responsiveness. In what follows, we detail the design of these two experiments.

3.1 Conjoint Experiment

Conjoint analysis is an experimental technique to solicit aggregate preferences through forced choices. In a conjoint experiment respondents are asked to decide between two or more hypothetical choices that vary across multiple attributes. Respondents make decisions without having to provide explanations, allowing their decisions to capture even politically or socially undesirable preferences that may not be explicitly expressed. Because the values for each attribute are randomly assigned, researchers may still estimate respondents' average preferences regarding each of the attributes by estimating their effects on conjoint decisions (Hainmueller, Hopkins and Yamamoto, 2014).

Our conjoint experiment presents survey respondents with a pair of hypothetical property tax policies. The policies are represented by four attributes, each of which captures a specific feature of a tax policy and can take one of two possible values as illustrated in Table 1.³⁴ Central to our research questions are the first two attributes. The first captures responsiveness—whether channels exist for citizen input on the use of property tax revenues. The second attribute reflects accountability—whether higher-level governments audit the allocation of property tax revenues and punish officials who embezzle from this fund. The remaining two attributes represent economic factors that determine tax compliance, namely the economic costs and benefits of the tax regime for individual citizens. Following standard practice for creating conjoint profiles, the order of the attributes was randomized for each

³⁴See Appendix A.3 for a screenshot of the conjoint experiment in Chinese presented to the respondents.

respondent; and the values of each attribute were randomly assigned in each round.³⁵

Tax Policy Attributes	Possible Values
Channels for citizen input on the use of tax revenue	YES or NO
Provincial government monitor and the use of tax revenue and punishment for corruption	YES or NO
Fines for late payment of taxation	15% or 3% of tax payment
Tax exemption on first apartment	YES or NO

Table 1: Conjoint experiment attributes and possible attribute values

Respondents participating in our conjoint experiment faced four rounds of forced decisions. In each round they were given two property tax designs and asked to decide between them. After seeing the two property tax designs and their attributes, respondents were asked to indicate which of the two designs they preferred and then scored their willingness to comply with each design on a 5-point Likert scale. Because the attribute values were randomized, we could estimate the causal effect of each attribute’s values on respondents’ decisions by regressing their answers on binary indicators representing each attribute value. These Average Marginal Component Effects (AMCEs) captured their preferences among the values of each attribute.

We embedded this conjoint experiment in a larger online survey of Chinese citizens. To avoid biases in sampling strategy, we commissioned two survey firms, Qualtrics and Survey Sampling International (SSI),³⁶ to separately implement this survey in their respondent pools. For both respondent pools, we used a quota sampling strategy based on age and gender and targeted the adult urban population.³⁷ Qualtrics and SSI successfully collected data from 464 and 431 respondents, respectively, between May 26 and June 6, 2016.³⁸

³⁵To ease the cognitive load on participants in our online conjoint experiment, we conducted a training session before the experiment, acquainting them with the format. Following the training, we administered a quiz to filter out individuals who could not grasp the format.

³⁶SSI later rebranded as Dynata.

³⁷The breakdown of our quotas is as follows: aged 25–34: 23%; aged 35–44: 25%; aged 45–54: 22%; aged 55+: 30%; men: 51%; women: 49%. These quotas were based on the 2010 China Census.

³⁸See Appendix A.1 for summary statistics of respondent characteristics. Appendix A.2 reports a balance

3.2 Lab-in-the-field Experiment

To create a simulated property tax environment, we adapted an iterated public goods game in which participants made decisions regarding compliance with property tax payments that fund local public goods and services. The main objective of the lab-in-the-field experiment was to address the concern that the conjoint experiment reflects only the institutional design of government responsiveness and accountability, but both could be unrealistic in the Chinese context. To address this concern, the experimental manipulation in the lab-in-the-field experiment was focused on whether subjects experience positive or negative responsiveness and accountability behavior by the government. We implemented the game in sessions with subjects randomly divided into four groups of approximately equal sizes. Each subject was given an iPad with a preinstalled app to interface with the game. We then provided instructions and conducted some training for every participant in a session.³⁹

Following the training we introduced the subjects to the simulation. We informed them that within the simulated property tax environment they would be required to pay property taxes annually, but they would be able to decide whether to comply with this requirement and pay the tax promptly or disobey it and delay payment.⁴⁰ In each round the property tax rate was fixed at 0.3%, but the property tax payment varied across subjects, depending on their property values. If subjects chose not to pay their property taxes on time, they faced a probability, the exact value of which was unknown to them, that they could be audited by the government, resulting in a penalty raising their payment to a value of 150% of their property tax. After our subjects had independently and anonymously made their decisions,

test verifying that the randomization of attributes was successful.

³⁹Our training detailed the development of property tax policy in China. We also discussed the common international practices of using property tax revenue to finance local public goods and services.

⁴⁰We framed the tax compliance decision as whether to pay property tax on time instead of whether to pay property tax at all. Based on our pilot study and focus groups, framing tax avoidance in the lab experiment would have been unrealistic for our subjects, who believe that complete avoidance of paying property tax is impossible; however, delaying property tax payments is possible because delayed payments could be passed over to future potential buyers, effectively allowing property owners to avoid paying. This practice is common in Shanghai.

we revealed the overall property tax compliance of the entire group.⁴¹ Finally, we informed subjects that all the property tax payments as well as penalties for late payments (if any) would go to the general government property tax revenue. Eighty percent of the revenues were then divided equally among all the subjects within the group as a lump sum payment representing a public good, the remaining 20% percent allocated to government overhead and waste.⁴²

To ensure the experiment was incentive-compatible, we tied their experimental compensation to their tax compliance decisions. All subjects were endowed with 15 units of token money at the beginning of the session. The amount of token money could vary after each round, depending on the size of their tax payment, potential penalty for noncompliance, and other subjects' tax compliance decisions. At the end of the experiment, we paid each subject an amount proportionate to their remaining token money, with amounts ranging between 180 and 300 RMBs, approximately the median daily salary in Shanghai.

In total, our subjects played 12 rounds of this game. Rounds 1–6 took place without any experimental manipulation. These six rounds allowed our subjects to familiarize themselves with the game settings. More importantly, we used the behavioral outcomes in these six rounds to establish a benchmark property tax compliance rate for each subject and to estimate the pace at which they adapted their strategies to information about other players' actions. These were necessary to help identify within-subject changes due to the experimental manipulation.

We introduced an experimental manipulation for the treatments of government responsiveness and accountability at the beginning of Round 7 and continuously reexposed respondents to it in all the remaining rounds. Each subject was randomly exposed to one of four possible treatment conditions, with all respondents in the same game session receiving the

⁴¹Specifically, the subjects were informed about the number of group members who paid taxes on time, the number of noncomplying group members who were caught by government auditing, and the total property tax revenues.

⁴²We informed the subjects that not all government revenues would be spent on public goods because of government operation costs and waste.

same treatment throughout Rounds 7–12. Figure 1 below illustrates the sequence of events in each of these rounds.

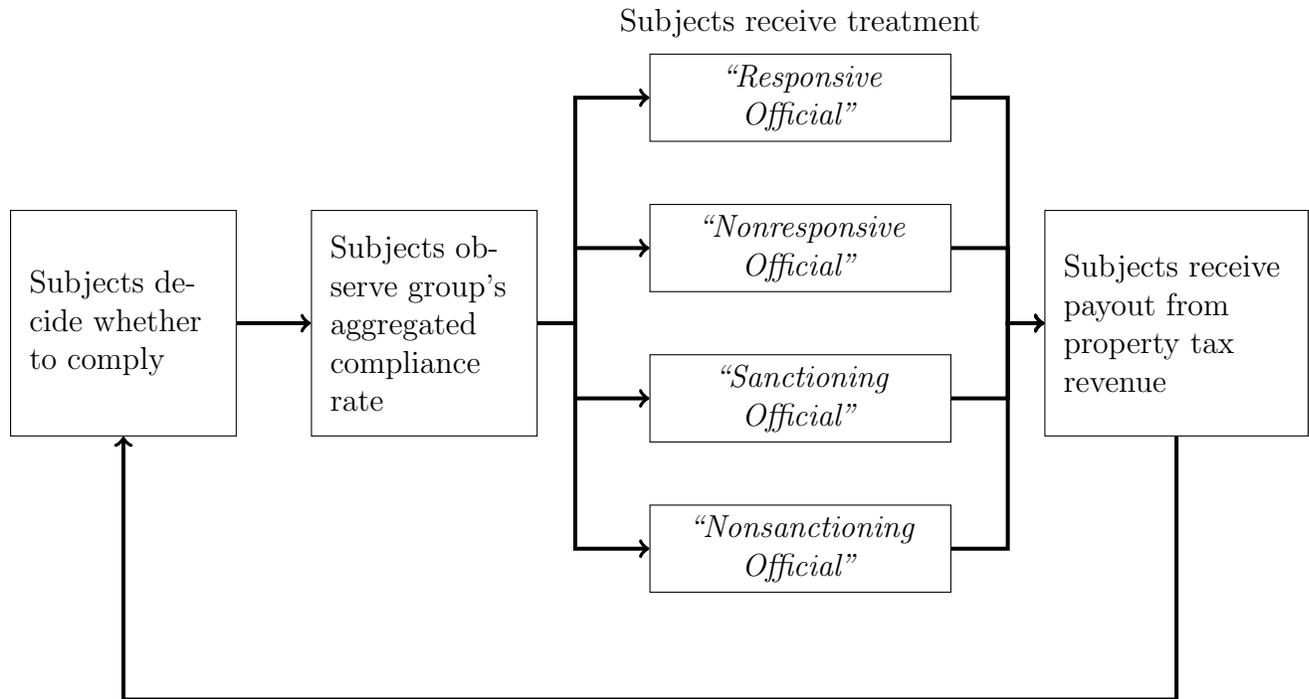


Figure 1: The step-by-step procedure for each round of the public goods game.

Among the four treatment conditions the first two signaled the degree of government responsiveness to citizen input. Under the first treatment condition, which we term “responsive official,” the subjects was given an opportunity to vote for their most preferred use of property tax revenue;⁴³ and they were informed, using animated videos designed to garner attention, that the government official in charge of this expenditure had respected the group’s vote outcomes. Under the second condition, which we termed “nonresponsive official,” the authority simply disregarded the group’s votes. These two treatment conditions reflected the level of responsiveness in many developing countries, where governments may not consistently address citizens’ demands even when formal channels for participation are in place.

The third and fourth treatment conditions pertain to top-down accountability. Under the

⁴³The choices were education, transportation, and community facilities.

third treatment condition, which we called “sanctioning official,” the subjects were informed that a corrupt official had stolen money from property tax revenue but had been caught and punished by the higher-level government. The fourth treatment condition, which we called “nonsanctioning official,” conveyed the opposite message: the corrupted official had not been punished. We adopted this anticorruption framing because the anticorruption campaign is a salient issue in China, where a large number of central and local government officials have been prosecuted for corruption charges since Xi Jinping came to power in 2012.

We conducted 16 lab-in-the-field experimental sessions in July 2018 in Shanghai, chosen as the experimental site for three compelling reasons. First, Shanghai is one of the two Chinese cities that had introduced property tax policies in 2011, making annual property tax payments a familiar concept to Shanghai residents compared to those in other regions of China. Second, Shanghai ranks among China’s most developed cities, boasting one of the country’s priciest real estate markets. Consequently, property tax is a prominent issue for Shanghai residents, who may have to bear a substantial tax burden. Last, Shanghai residents generally have higher levels of education, ensuring a better understanding of the rules governing our behavioral experiment. To closely replicate real-world conditions, we enlisted participants from the general population, excluding full-time students. Subjects were recruited through advertisements on social media. In total, we collected lab experiment data from 123 subjects who are homeowners,⁴⁴ summary statistics for whom are detailed in Table C-4.⁴⁵

⁴⁴Our sample size was relatively small, reflecting the financial and logistical challenges in recruiting subjects to participate in an hour-long in-person simulation. We attempt to mitigate the lack of statistical power by leveraging repeated measures from our respondents as they completed the 12 rounds of the game. Section 4.2 details our analytical approach.

⁴⁵During the experimental session, we identified a few subjects whose data might be subject to various biases. For example, some older subjects had problems understanding the rules of our property tax compliance game and required our student assistants to help interpret the rules throughout the entire lab session. Hence, we excluded a total of five such respondents in our analyses.

4 Results

4.1 Conjoint Experiment

Figure 2 reports estimates of the AMCEs from the baseline model along with 95% confidence intervals. We present results from the SSI sample (as triangles), the Qualtrics sample (squares), and combined results from both samples (circles). As shown, all our AMCE estimates are broadly consistent across both samples; therefore, we focus on results from the combined sample in the remaining part of the section, leveraging the efficiency benefit from the larger sample size.

Notably, each point estimate captures the effect of each attribute on the probability that a respondent would prefer a tax policy. Because each attribute takes one of two possible values, we designate one of the values as the baseline. Intuitively, the AMCE captures how much greater the likelihood that a respondent would prefer a particular tax policy with a non-baseline value for an attribute over another with the baseline value for the same attribute is, all else being equal. For example, the AMCEs for the “Tax Payment Penalty” attribute indicate how often respondents preferred a tax policy with a 3% penalty over another with a 15% penalty.

Our results show that Chinese respondents cared about both responsiveness and accountability. The AMCEs of the corresponding attributes—bottom-up channels for citizen inputs and top-down monitoring and sanctioning—were both estimated to be substantively and statistically significant. Specifically, we found a property tax policy with channels for citizen input about the use of tax revenue 12.4 percentage points more likely to be preferred over a tax policy without them ($CI = [9.9, 14.9]$). Meanwhile, the inclusion of avenues for top-down monitoring and sanctioning was found to increase support for a tax policy by 9.8 percentage points ($CI = [7.4, 12.1]$).

These estimates are comparable to the AMCE estimate for tax exemption in terms of magnitude. Specifically, an exemption for a first apartment increased support for a property

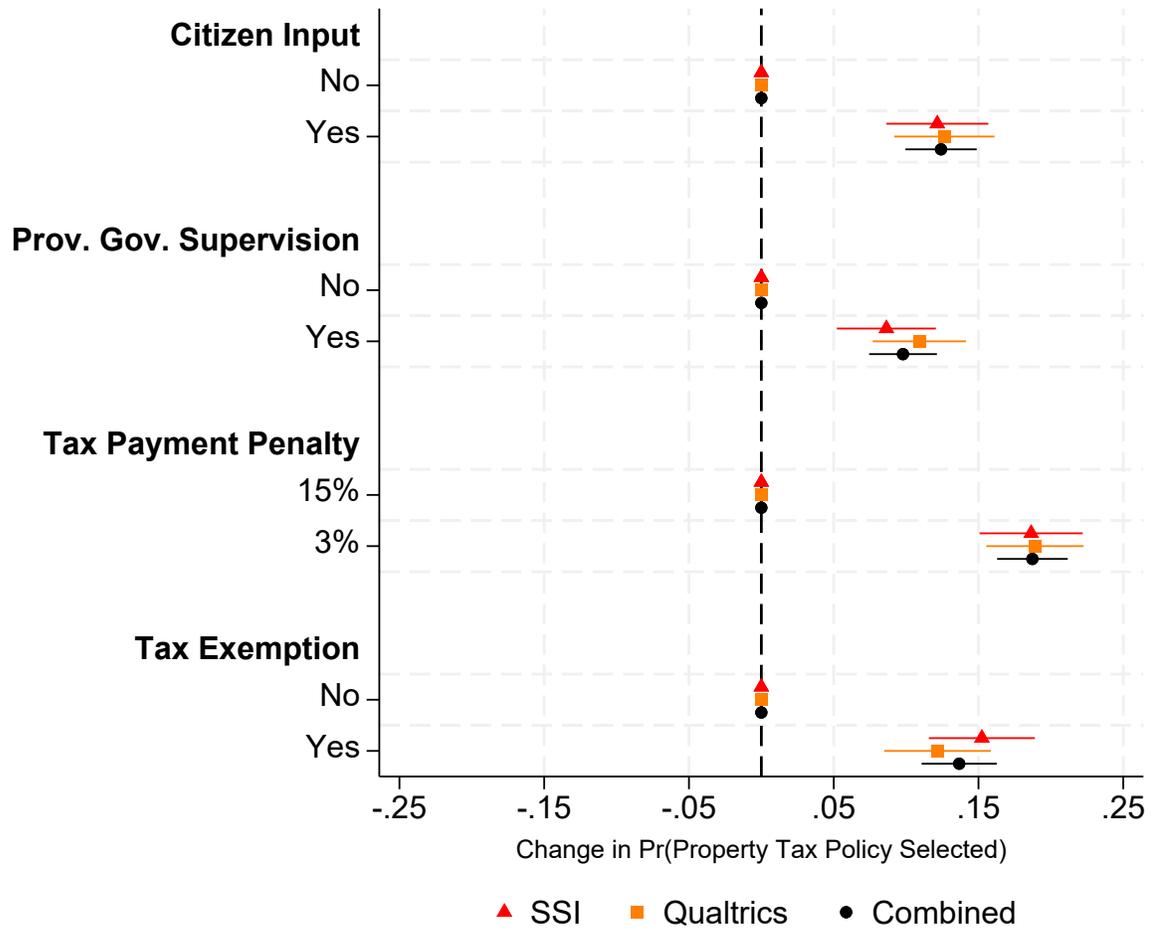


Figure 2: Estimated AMCEs of conjoint attributes on tax policy preference

tax policy by 13.7 percentage points ($CI = [11.0, 16.2]$). Given the preoccupation with the scope of exemptions in the existing debate, our finding suggests that accountability and responsiveness matter just as much as exemptions and deserves at least an equal amount of attention. Finally, consistent with economic models of tax compliance, we found that a penalty of 3% for non-payment increased support for a tax policy by 18.7 percentage points ($CI = [16.2, 21.2]$) compared to a harsher penalty of 15%.

We conducted additional analyses to probe the robustness of our main findings. Our first concern regarded external validity, given that property ownership was not universal among our respondents and most Chinese residents were still unfamiliar with the practice of property taxation. Moreover, property tax is a form of redistribution from property owners to nonowners; thus these two groups of individuals may have had distinct preferences for property tax. To address this concern, we conducted a subset analysis of our data, separating our respondents by the number of apartments they own. As shown in Figure B-3 in the Appendix, these additional analyses still prove consistent with the main results. Specifically, we find similar preferences experienced by those who owned no apartments, those who owned only one apartment—likely their primary home, and those who owned multiple apartments, because the estimates among these different groups of subjects were statistically indistinguishable.

We also evaluated the concern that direct experiences with actual property taxation might influence preferences for property tax policy. Two cities—Shanghai and Chongqing—have had pilot property tax regimes in place since 2011. We thus examined the results for residents living there versus residents of all other cities. Our analysis does not show markedly different results between residents with direct experience with the property tax and those without experience (Figure B-4 in the Appendix). The estimates for top-down institutions for sanctioning and bottom-up channels for citizen input are broadly compatible in the two subsets of the respondents.

Finally, we evaluated the respondents' stated willingness to comply with property tax

as alternative outcome of interest. Figure B-2 in Appendix B.1 reports the baseline results, and they are broadly consistent with the findings above on tax policy preferences. In this analysis, the size of the effects of channels for citizen input about the use of tax revenue and institutions for auditing and punishing misuse of tax revenue roughly align with the size of the effect of tax exemptions. Surprisingly, we found little evidence that low penalties increase tax compliance.

4.2 Lab-in-the-Field Experiment

Turning to the analysis of the lab-in-the-field experiment, Figure 3 shows the average tax compliance rate by experiment round for respondents in each of the four treatment conditions. Because subjects were exposed to the treatment videos after making their compliance decision in Round 7, Rounds 1–7 constitute the pre-treatment period, and Rounds 8–12 constitute the post-treatment period. The bold horizontal lines in Figure 3 indicate the average pre- and post-treatment compliance rates.

Initially, the average property tax compliance rates ranged from 65% to 73% before our experimental manipulation, indicating a notably high degree of property tax compliance among our participants. This finding sharply contrasts with the low compliance rates typically observed in many developing countries⁴⁶ largely because China has stronger bureaucratic and coercive capacities than many other developing countries. Moreover, these compliance rates align with a national survey conducted by other researchers one year later, which shows that 68% of urban households did not oppose the introduction of property tax (Liu, Gao and Tao, 2019).

Looking at participants’ behaviors after the introduction of experimental manipulations, we found strong evidence that the “sanctioning official” treatment improved tax compliance but no evidence for the “responsiveness official” treatment. Most succinctly, Figure 3 plots the average tax compliance rate by experiment round for respondents in each of

⁴⁶For instance, Weigel (2020)’s experimental treatment increased property compliance rate from 0% to 10% in Congo.

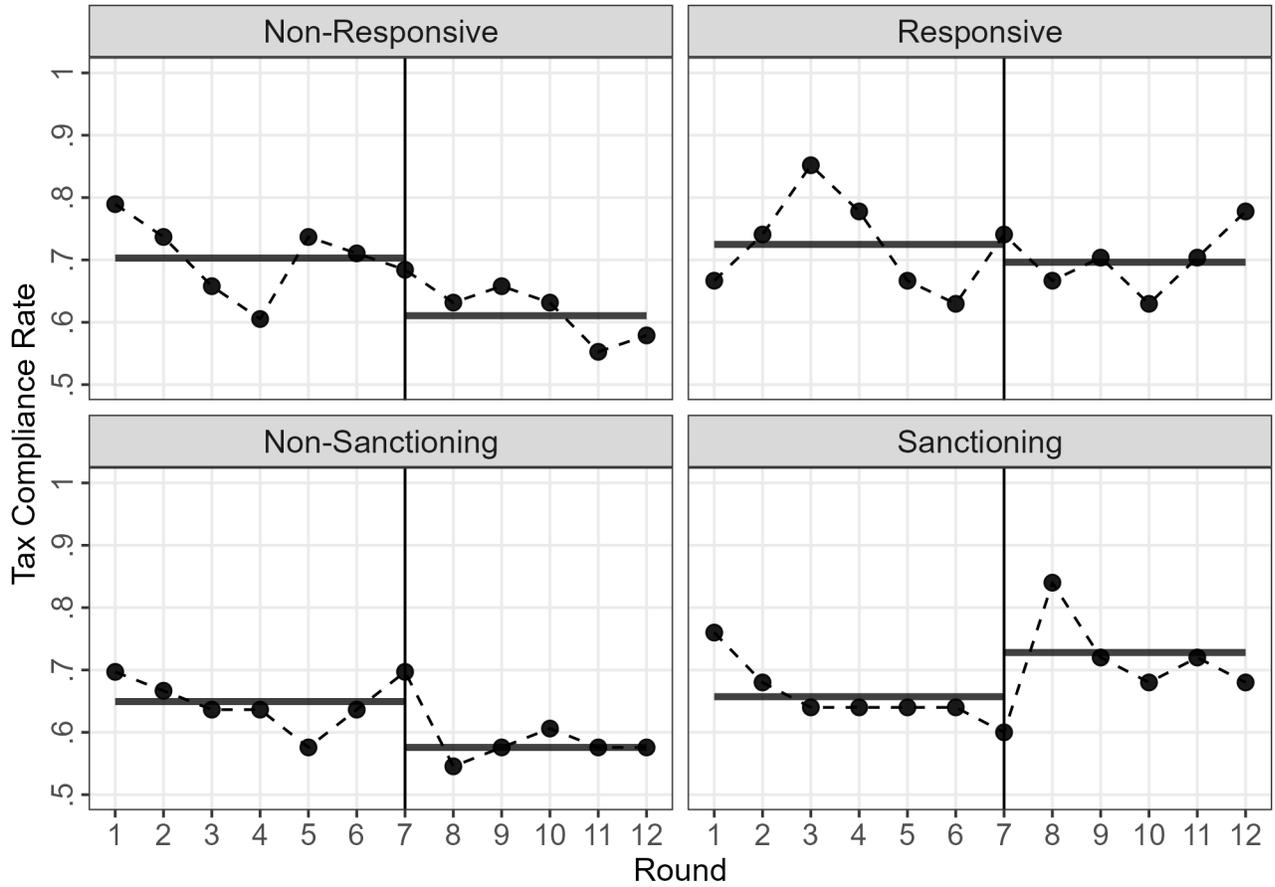
the four treatment conditions. Among the four treatment groups, only in the “sanctioning official” treatment group did respondents’ tax compliance rate increase following exposure to treatment. In contrast, compared to the pre-treatment level, post-treatment compliance rates appear markedly lower in the “nonresponsive official” and the “nonsanctioning official” treatment groups and slightly lower in the “responsive official” treatment group.

Although evidence from Figure 3 strongly suggests a positive effect of the “sanctioning official” treatment on tax morale, it also reveals complications that impede a straightforward conclusion. Most notably, in all four treatment groups, tax compliance appears to decrease between rounds, even in the absence of treatment exposure. Although unforeseen in our survey design process, this phenomenon is not surprising. As respondents proceeded through the 12 rounds of the experiment, fatigue and boredom may have set in, potentially causing a gradual increase in noncompliance behavior. More importantly, respondents may have noticed that noncompliance is only occasionally punished, and as they acquired more experience with the experiment and learned to act strategically, they reduced compliance in later rounds. Because our videos appeared in the second half of the experiment, treatment exposure directly correlates with round numbers, allowing the decrease in compliance over time to confound treatment effect estimates. Thus it is imperative for us to account for time trends in the analysis of experimental data.

To address the concerns outlined above, Table 2 presents results from regression analyses in an attempt to account for time trends as well as improve estimation precision through statistical controls. It shows coefficient estimates from regressions of treatment group indicators on respondents’ tax compliance. For each respondent-round, we coded four treatment indicators, each of which is a dummy variable equaling one if the respondent was in the respective treatment condition and the round was a post-treatment. Tax compliance is a binary indicator, allowing each coefficient estimate to be interpreted as the effect of the respective treatment video on the expected compliance rate.

Models 1 and 2 in Table 2 present results from analyses that make use of responses only

Tax Compliance Rate by Experiment Round



Note: Tax compliance rate by experiment round for respondents in each treatment condition. Bold horizontal lines denote average tax compliance rates before (Rounds 1–7) and after (Rounds 8–12) respondents were exposed to the treatment.

Figure 3: Tax compliance rate by experiment round for respondents in four treatment conditions.

Table 2: Impact of different types of government officials on tax compliance

Dependent Variable: Model:	Tax Compliance				
	Immediate Effect		Average Effect		
	(1)	(2)	(3)	(4)	(5)
<i>Variables</i>					
Non-Responsive	-0.027 (0.077)	-0.079 (0.069)	-0.083* (0.042)	-0.043 (0.045)	-0.039 (0.062)
Responsive	0.008 (0.086)	0.037 (0.065)	-0.019 (0.027)	-0.003 (0.043)	-0.005 (0.060)
Non-Sanctioning	-0.113 (0.085)	-0.091 (0.080)	-0.045 (0.047)	-0.031 (0.059)	-0.087 (0.067)
Sanctioning	0.181** (0.084)	0.200* (0.115)	0.040 (0.056)	0.123** (0.057)	0.192** (0.081)
<i>Fixed-effects</i>					
Respondent		Yes	Yes	Yes	Yes
<i>Time trends</i>					
Common				Yes	
Individual					Yes
<i>Fit statistics</i>					
Observations	246	246	1,476	1,353	1,353
R ²	0.023	0.779	0.542	0.552	0.634

Clustered (Respondent) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

from rounds 6 and 8 of the experiment. Because the treatment video appeared in round 7, rounds 6 and 8 are the pre- and post-treatment rounds closest to the time of exposure; a comparison of tax compliance between these two rounds is less likely to be contaminated by over-time trends. We call the effect identified by this comparison the “immediate effect.” Model 1 is a standard OLS regression of tax compliance on four treatment indicators.⁴⁷ In Model 2 we added respondent fixed effects to absorb all idiosyncratic variations across individuals and improve precision; because treatment assignment is randomized, we are not concerned with bias. Still, because they estimate four different treatment effects using only two data points per each of our 123 respondents, both Models 1 and 2 are statistically underpowered, with Model 2 offering only some minor improvements.

To ameliorate the power problem, Models 3, 4, and 5 estimate the average effect over the entire duration of the experiment instead of the immediate effect as in Models 1 and 2. Model 3 includes respondent fixed effects but does not attempt to control for time trends. When no time trends exist, this model would offer the greatest statistical power thanks to repeated measures for each respondent; however, when time trends exist, as in the case of our experiment, it would lead to bias. We present its results for transparency even though it is not our preferred model.

Model 4, our preferred model, controls for a common linear time trend by including the round number as a covariate. When there is no time trends, Model 4 lags behind Model 3 in terms of statistical power, but this difference shrinks under the presence of time trends. In Model 5 we took a step further by using individual time trends through respondent-specific time slopes.

Table 2 indicates a clear, consistent, and positive effect of the “sanctioning official” treatment on tax compliance. Except for Model 3, the effect is statistically significant throughout, with magnitudes indicating an increase in tax compliance rates from 12.3 percentage points (Model 4) to 20.0 percentage points (Model 2). In contrast, none of the coefficient estimates

⁴⁷We also estimate an intercept term but omit it from Table 2.

on the remaining treatment effect indicators are statistically significant in a consistent manner. If anything, the effect estimates that most consistently follow in terms of magnitude pertain to the “nonsanctioning official” treatment. These estimates consistently show a negative trend, which, when juxtaposed with the positive estimates for the “sanctioning official” treatment, underscore the stark contrast between upholding top-down sanctioning and its absence. This distinction, however, is not as evident in the case of the “nonresponsive official” and “responsive official” treatments.

Notably, our analyses draw from a relatively small sample of 123 respondents divided over four treatment groups, resulting in an inevitable statistical power problem. The use of repeated measures for each respondent partially mitigates this problem, but ultimately even our most efficient analytical approaches are still powered to detect only very large treatment effects. Specifically, as shown in our power analysis in Appendix C.9, under ideal conditions where no time trends exist, depending on the extent of heterogeneity across respondents, Model 3 may have more than 90% power to detect an effect of 20 percentage points, between 70–90% power to detect an effect of 15 percentage points, but only 40–60% power to detect an effect of 10 percentage points. Under more realistic conditions when a negative time trend exists, Model 4, our preferred model, has more than 90% power to detect an effect of 20 percentage points, between 70–75% power to detect an effect of 15 percentage points, but only 30–55% power to detect an effect of 10 percentage points. Given a pre-treatment compliance rate of 70%, even a 10 percentage point effect is already a very sizeable effect. Importantly, however, our treatment effect estimates for the “sanctioning official” condition consistently fall above this threshold while estimates for the three remaining conditions are much closer to zero, offering assurance that the statistical (in)significance of these estimates is not a mere accident.

We also conducted several additional analyses to evaluate the robustness of our findings and present them in Appendix C. First, we conducted analyses by separating the samples by apartment ownership characteristics (e.g., number of apartments owned, apartment values).

Because of a smaller sample size, we found only suggestive evidence that the estimate of the sanctioning treatment is larger with greater home ownership but also with lower property values (Table C-9). In addition, we found that our results are robust to dropping the first and last rounds (Table C-5) as well as to dropping the seventh round (Table C-6). The reason to exclude these rounds is that the first and last rounds can introduce confounding because respondents may see the first as a practice round—thus not investing sufficient effort in their tax compliance decisions—and expect the last round to be consequence-free, thus behaving too strategically. The seventh round is a potential source of noise because it is the round in which respondents were introduced to the treatment. We also attempted to replace individual fixed effects with session fixed effects (Table C-7) and replace the linear time trends with higher-order polynomials (Table C-8). All these robustness checks yielded results substantively identical to our main analysis.

Finally, we assessed the external validity of our lab experiment results. In particular, we requested the subjects in our lab experiment to complete a postexperiment survey, which included a similar conjoint experiment that we initially designed for the nationwide survey conducted over the Internet. As reported in Figure C-6 in the Appendix, the conjoint experiment reveals strikingly similar results compared to the main results based on the online conjoint experiment reported in Figure 2.

5 Conclusion

Property tax has been a prominent source of revenue for the state, but the adoption of this tax typically produce impasses, both administrative and political. In this paper we delved into the impact of institutions for accountability and responsiveness in enhancing property tax compliance. We conducted both conjoint experiments and lab-in-the-field experiments in urban China. Our conjoint experiment provides supporting evidence that Chinese citizens value both types of institutions, enhancing their support of a property tax policy. The

marginal effects of these two attributes are comparable to factors in the classic economic model of tax evasion. Our lab-in-the-field experiment reveals more intriguing results on tax compliance behaviors. Specifically, we have found that experiencing top-down accountability is more important in enhancing tax compliance than experiencing government responsiveness to citizen input.

Notably, our findings do not suggest that economic considerations are inconsequential or peripheral to Chinese citizens' support for property taxes. Using simulations based on survey data, some scholars have concluded that a tax rate of 0.5% to 1.0% is likely to be more acceptable to urban homeowners.⁴⁸ Similarly, offering property tax exemptions under certain criteria is another widely supported proposition among urban residents.⁴⁹ In addition, fiscal transparency in the use of tax revenue is a crucial factor in garnering citizen support and ensuring compliance with tax policies, a point reinforced by the findings of our study.⁵⁰ Finally, citizen support for property taxes is inherently tied to the broader tax regime and economic conditions that they experience. Although the Chinese government imposes minimal taxes on personal income and wealth, social security contributions continue to represent a substantial deduction from individuals' earnings. Finally, any reform that increases the tax burden on citizens or firms is likely to trigger strong resentment and encourage tax avoidance, particularly during economic downturns.⁵¹

Subject to these conditions of the broader context of property taxation in China, our study generates several important implications for understanding the relationship among representation, legitimacy, and tax compliance. Institutions for citizen participation in shaping government decision-making in nondemocratic regimes like China have proved effective

⁴⁸Cao and Hu (2016) and Zhang and Hou (2016).

⁴⁹Hou (2018) and Zhang and Hou (2016).

⁵⁰See Luttmer and Singhal (2014) for a comprehensive review of the relationship between government transparency and tax compliance, particularly in studies utilizing field experiments. Kao, Lü and Queralt (2024) offer direct experimental evidence on fiscal transparency and tax compliance in China.

⁵¹Zhang and Jing (2020) discuss the strategies and challenges in property tax reform in China. Cheng, Chen and Qi (2023) demonstrate that firms underreported their profit margins to evade taxes following China's 2008 corporate tax reform.

in some instances but lacking in others.⁵² Promoting accountability or responsiveness has very distinct implications for institution building. On one hand, signaling responsiveness emphasizes the role of ordinary citizens during the policymaking process, thus requiring building institutions that enhance grassroots political participation and government responsiveness. Signaling accountability, on the other hand, could be achieved by strengthening top-down bureaucratic sanctioning without devolving political power to ordinary citizens. To this end, it is important to determine whether strengthening the institutions to signal bottom-up responsiveness, or projecting top-down accountability through bureaucratic sanctioning without distributing political power to ordinary citizens, would generate greater policy compliance.

We provide experimental evidence from a nondemocratic context, demonstrating that promoting responsiveness to citizen demands can legitimize government policies; but this bottom-up approach may prove ineffective in states with weak or absent democratic institutions. The ineffectiveness of citizen engagement in the developing world is not limited to tax compliance. A growing number of experimental studies have shown that providing information on government performance did not engender greater political participation to enhance accountability and local governance.⁵³ In addition, participatory budgeting has not consistently delivered its promise to enhance public goods provision.⁵⁴ Finally, increasing grassroots participation in monitoring officials could have little impact on reducing corruption (Olken, 2007).

Our findings suggest that the components of accountability must be unbundled and that principals can rely on third parties to take responsibility for some or all of these components. Accountability requires institutions that enable principals to establish the duties of their agents, monitor the behavior of their agents, and sanction them for misbehavior. Ac-

⁵²For rural elections, see O'Brien and Han (2009). For election of local people's congress, see Manion (2015). Qiaoan and Teets (2020) offer a review of government responsiveness under the leadership of Hu Jintao and Xi Jinping.

⁵³Kosack and Fung (2014) review the mixed evidence of the impact of transparency on governance.

⁵⁴See Speer (2012) for a review of the limitations of participatory budgeting.

countability sometimes entails multiple principals working together to hold the same set of agents accountable (Braithwaite, 1998). The role of one set of principals (citizens) may be to set policy priorities or monitor the performance of local officials, but another set of principals (higher-level officials) sanctions local officials for poor performance or for failing to respond to the priorities set by citizens. Together, these components of accountability may complement or compensate for one another, ultimately enhancing citizen tax compliance.

References

- Ali, Merima, Odd-Helge Fjeldstad and Ingrid Hoem Sjørusen. 2014. "To Pay or Not to Pay? Citizens' Attitudes Toward Taxation in Kenya, Tanzania, Uganda, and South Africa." *World Development* 64:828–842.
- Allingham, Michael G. and Agnar Sandmo. 1972. "Income tax evasion: a theoretical analysis." *Journal of Public Economics* 1(3):323–338.
- Antinyan, Armenak, Zareh Asatryan, Zhixin Dai and Kezhi Wang. 2021. "Does the frequency of reminders matter for their effectiveness? A randomized controlled trial." *Journal of Economic Behavior & Organization* 191:752–764.
- Bates, Robert H. and Da-Hsiang Lien. 1985. "A Note on Taxation, Development, and Representative Government." *Politics & Society* 14(1):53–70.
- Bernstein, Thomas P. and Xiaobo Lü. 2003. *Taxation without Representation in Rural China*. New York: Cambridge University Press.
- Bianco, Lucien. 2001. *Peasants without the party : grass-roots movements in twentieth-century China*. Asia and the Pacific Armonk, N.Y.: M.E. Sharpe. 00052211 Lucien Bianco. maps ; 24 cm. "An East Gate Book." Includes bibliographical references and index. Asia and the Pacific (Armonk, N.Y.).
- Blau, Peter M. 1964. *Exchange and power in social life*. New York,: J. Wiley.
- Bodea, Cristina and Adrienne LeBas. 2014. "The Origins of Voluntary Compliance: Attitudes toward Taxation in Urban Nigeria." *British Journal of Political Science* 46(1):215–238.
- Braithwaite, John. 1998. Institutionalizing Distrust, Enculturating Trust. In *Trust and Governance*, ed. Valerie Braithwaite Levi and Margaret. New York: Russell Sate pp. 343–375.
- Cao, Jing and Wenhao Hu. 2016. "A microsimulation of property tax policy in China." *Journal of Housing Economics* 33:128–142.
- Carlsmith, Kevin M and John M Darley. 2008. "Psychological aspects of retributive justice." *Advances in experimental social psychology* 40:193–236.
- Castañeda, Néstor, David Doyle and Cassilde Schwartz. 2020. "Opting Out of the Social Contract: Tax Morale and Evasion." *Comparative Political Studies* 53(7):1175–1219.
- Chen, Jidong, Jennifer Pan and Yiqing Xu. 2016. "Sources of Authoritarian Responsiveness: A Field Experiment in China." *American Journal of Political Science* 60(2):383–400.
- Cheng, Hua, Xiaowei Chen and Shusen Qi. 2023. "Asymmetric corporate tax compliance: Evidence from a tax reform in China." *China Economic Review* 79.
- Cohen-Charash, Yochi and Paul E. Spector. 2001. "The Role of Justice in Organizations: A Meta-Analysis." *Organizational Behavior and Human Decision Processes* 86(2):278–321.
- Collier, Paul, Edward Glaeser, Tony Venables, Priya Manwaring and Michael Blake. 2017. "Land and property taxes: exploiting untapped municipal revenues." *Policy brief* .

- Cui, Wei. 2022. *The Administrative Foundations of the Chinese Fiscal State*. Cambridge: Cambridge University Press.
- Dahl, Robert A. 1956. *A preface to democratic theory*. Charles R Walgreen Foundation lectures Chicago: University of Chicago Press. 56006642 illus. 21 cm.
- De Cremer, David and Tom R. Tyler. 2005. Managing Group Behavior: The Interplay Between Procedural Justice, Sense of Self, and Cooperation. In *Advances in Experimental Social Psychology*. Vol. Volume 37 Academic Press pp. 151–218.
- de la Cuesta, Brandon, Lucy Martin, Helen V. Milner and Daniel L. Nielson. 2022. “Owning It: Accountability and Citizens’ Ownership over Oil, Aid, and Taxes.” *The Journal of Politics* 84(1):304–320.
- de la Cuesta, Brandon, Lucy Martin, Helen V. Milner and Daniel L. Nielson. 2023. “Do Indirect Taxes Bite? How Hiding Taxes Erases Accountability Demands from Citizens.” *The Journal of Politics* 85(4):1305–1320.
- Distelhorst, Greg and Yue Hou. 2017. “Constituency Service under Nondemocratic Rule: Evidence from China.” *The Journal of Politics* 79(3):1024–1040.
- Fearon, James, D. 1999. *Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance*. Cambridge: Cambridge University Press pp. 55–97.
- Field, Lars P. and Bruno S. Frey. 2007. “Tax Compliance as the Result of a Psychological Tax Contract: The Role of Incentives and Responsive Regulation.” *Law & Policy* 29(1):102–120.
- Flores-Macías, Gustavo A. 2018. “Building support for taxation in developing countries: Experimental evidence from Mexico.” *World Development* 105:13–24.
- Göbel, Christian. 2010. *The politics of rural reform in China : state policy and village predicament in the early 2000s*. Chinese worlds Milton Park, Abingdon, Oxon ; New York, NY: Routledge.
- Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2014. “Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments.” *Political Analysis* 22(1):1–30.
- Hou, Yilin. 2018. *Development, governance, and real property tax in China*. Springer. 2018950843 Yilin Hou.
- Hou, Yilin, Qiang Ren and Ping Zhang. 2014. *The property tax in China: History, pilots, and prospects*. Vol. 1 Springer.
- Kao, Jay C., Xiaobo Lü and Didac Queralt. 2024. “Do Gains in Political Representation Sweeten Tax Reform in China? It Depends on Who You Ask.” *Political Science Research and Methods* 12(1):146–165.
- Khanna, Jyoti and Michael Johnston. 2007. “India’s middlemen: connecting by corrupting?” *Crime, Law and Social Change* 48(3):151–168.
- Kosack, Stephen and Archon Fung. 2014. “Does Transparency Improve Governance?” *Annual Review of Political Science* 17(1):65–87.

- Lardy, Nicholas R. 2019. *The state strikes back: the end of economic reform in China?* Peterson Institute for International Economics.
- Levi, Margaret. 1989. *Of rule and revenue*. University of California Press.
- Levi, Margaret and Audrey Sacks. 2009. “Legitimizing beliefs: Sources and indicators.” *Regulation & Governance* 3(4):311–333.
- Levi, Margaret and Laura Stoker. 2000. “Political Trust and Trustworthiness.” *Annual Review of Political Science* 3(1):475–507.
- Li, Wen. 2014. “Wo guo fang di chan shui shou ru shu liang ce suan ji qi chong dang di fang shui zhu ti shui zhong de ke xing xing fen xi [On the estimation of the house property tax revenue and the feasibility of the house property tax acting as the main tax of the local tax system in China].” *Cai mao yan jiu [Finance and Trade Economics]* 9:14–25.
- Liu, Hua, Lipeng Chen and Juan Wei. 2020. “Fang di chan shui shou ru yong tu dui jun ming na shui yi yuan de ying xiang [Influence of the Purpose of Real Estate Tax Revenue on Residents’ Willingness to Pay the Tax].” *Cai mao yan jiu [Finance and Trade Economics]* 41(10):34–48.
- Liu, Jingdong, Fengqing Gao and Ran Tao. 2019. “Fang di chan shui de zhi fu yi yuan yu na shui neng li feng xi-ji yu 130 ge cheng shi de jia ting diao cha [Assessing Willingness and Capacibility to Pay Property Tax—A Household Survey based on 130 cities].” *Shui wu yan jiu [Taxation Research]* 8:51–58.
- Liu, Rong, Wei Zhang and Lingshuang Chen. 2015. “Fang di chan shui fei jian (shu) mian bi lv de gu ji yu qian zhan shui shou shou ru neng li de ce suan [On the evaluation and analysis of the total real estate tax-based on the data from China household finance survey].” *Cai mao yan jiu [Finance and Trade Economics]* 1:54–64.
- Luttmer, Erzo F. P. and Monica Singhal. 2014. “Tax Morale.” *Journal of Economic Perspectives* 28(4):149–68.
- Man, Joyce Yanyun. 2011. *China’s housing reform and outcomes*. Cambridge, MA: Lincoln Institute of Land Policy.
- Manion, Melanie. 2015. *Information for Autocrats: Representation in Chinese Local Congresses*. Cambridge Studies in Comparative Politics Cambridge: Cambridge University Press.
- Martin, Lucy. 2023. *Strategic Taxation: Fiscal Capacity and Accountability in African States*. Oxford University Press.
- Murphy, Kristina, Tom R. Tyler and Amy Curtis. 2009. “Nurturing regulatory compliance: Is procedural justice effective when people question the legitimacy of the law?” *Regulation & Governance* 3(1):1–26.
- Norregaard, Mr John. 2013. *Taxing immovable property revenue potential and implementation challenges*. International Monetary Fund.
- North, Douglass C. and Barry R. Weingast. 1989. “Constitutions and Commitment: The Evolution of Institutional Governing Public Choice in Seventeenth-Century England.” *The Journal of Economic History* 49(4):803–832.

- O'Brien, Kevin J. and Rongbin Han. 2009. "Path to Democracy? Assessing village elections in China." *Journal of Contemporary China* 18(60):359–378.
- Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115(2):200–249.
- Paler, Laura. 2013. "Keeping the Public Purse: An Experiment in Windfalls, Taxes, and the Incentives to Restrain Government." *American Political Science Review* 107(04):706–725.
- Pearson, Margaret M., Meg Rithmire and Kellee Tsai. 2023. *The State and Capitalism in China*. Cambridge University Press.
- Qiaoan, Runya and Jessica C Teets. 2020. "Responsive authoritarianism in China—a review of responsiveness in xi and Hu administrations." *Journal of Chinese political science* 25(1):139–153.
- Rothstein, Bo. 2009. "Creating political legitimacy: Electoral democracy versus quality of government." *American behavioral scientist* 53(3):311–330.
- Slemrod, Joel. 2019. "Tax Compliance and Enforcement." *Journal of Economic Literature* 57(4):904–54.
- Speer, Johanna. 2012. "Participatory Governance Reform: A Good Strategy for Increasing Government Responsiveness and Improving Public Services?" *World Development* 40(12):2379–2398.
- Timmons, Jeffrey F. and Francisco Garfias. 2015. "Revealed Corruption, Taxation, and Fiscal Accountability: Evidence from Brazil." *World Development* 70:13–27.
- Tsai, Lily L. 2021. *When People Want Punishment: Retributive Justice and the Puzzle of Authoritarian Popularity*. Cambridge University Press.
- Tsai, Lily L., Minh Trinh and Shiyao Liu. 2022. "What Makes Anticorruption Punishment Popular? Individual-Level Evidence from China." *The Journal of Politics* 84(1):602–606.
- Tyler, Tom R. 2006. *Why people obey the law*. Princeton: Princeton university press.
- Vigoda-Gadot, Eran. 2007. "Citizens' Perceptions of Politics and Ethics in Public Administration: A Five-Year National Study of Their Relationship to Satisfaction with Services, Trust in Governance, and Voice Orientations." *Journal of Public Administration Research and Theory* 17(2):285–305.
- von Hirsch, Andreas, Julian V Roberts, Anthony E Bottoms, Kent Roach and Mara Schiff. 2003. *Restorative justice and criminal justice: Competing or reconcilable paradigms*. Bloomsbury Publishing.
- Weigel, Jonathan L. 2020. "The Participation Dividend of Taxation: How Citizens in Congo Engage More with the State When it Tries to Tax Them*." *The Quarterly Journal of Economics* 135(4):1849–1903.
- Xu, Tong, Yue Yin and Jing Wang. 2023. "Wo guo jia ting de zhu fang xu qiu tai xing yu zui you fang chan shui she ji [The Housing Demand Elasticity and Optimal Property Tax Design in China]." *Jing ji yan jiu [Economic Research Journal]* 8.

- Yang, Yaowu and Yuansheng Liu. 2023. “Fang di chan shui de hong guan jing ji xiao ying yu fen pei xiao ying yan jiu [Research on the Macroeconomic Effects and Distributional Effects of Real Estate Tax Reform].” *Jing ji yan jiu [Economic Research Journal]* 5:41–59.
- Zhan, Peng and Shi Li. 2015. “Wo guo ju min fang chan shui yu shou ru bu ping deng [Property Tax and Income Inequality among Our Country’s Resident].” *Jing ji xue dong tai [Economic Perspectives]* 7:14–24.
- Zhang, Ping and Yijia Jing. 2020. “Strategies in Adopting Unpopular Policies in China: The Case of Property Tax Reform.” *Journal of Contemporary China* 29(123):387–399.
- Zhang, Ping and Yilin Hou. 2016. “Fang di chan shui de na shui neng li, shui fu feng bu ji zhan feng pei xiao ying [A Model for the Ability-to-pay Index of China’s Real Property Tax, Tax Burden Distribution, and Redistributive Effects].” *Jing ji yan jiu [Economic Research Journal]* 12:118–132.
- Zhang, Ping, Yilin Hou and Bo Li. 2020. “Fang di can shui yu fang jia yu zu jin-li lun mo ni ji qi dui zhong guo fang di chan shui kai zheng shi ji de qi shi [Impacts of Property Tax Levy on Housing Price and Rent: Theoretical Models and Simulation with Insights on the Timing of China Adopting the Property Tax].” *Cai mao yan jiu [Finance and Trade Economics]* 41(11):35–50.
- Zhang, Ping, Zhiyong Yao and Yinan Feng. 2021. “Fang di chan shui jian jing gai ge li jing–ji yu shui zhi yao su she ji de shi zheng yan jiu.” *Jing ji li lun yu jing ji guan li [Economic Theory and Business Management]* 12:79–92.
- Zhang, Yueran. 2020. “Political Competition and Two Modes of Taxing Private Homeownership: A Bourdieusian Analysis of the Contemporary Chinese State.” *Theory and Society* 49(4):669–707.
- Zhou, Chuang and Min Pan. 2021. “Fang di chan gai ge, jing ji zhen zhang yu jing rong wen ding [Property Tax Reform, Economic Growth, and the Financial Stability].” *Cai mao yan jiu [Finance and Trade Economics]* .

****NOT FOR PUBLICATION****

Supplementary Online Appendices

“Do Government Accountability and Responsiveness Enhance Support for Property Taxes? Experimental Evidence from China”

These appendices contain materials, results and robustness checks that supplement the main text.

A	Conjoint Experiment Implementation Details	2
A.1	Conjoint Experiment Subject Characteristics	2
A.2	Balance Test for the Conjoint Experiment	3
A.3	Conjoint Experiment Screenshot	5
B	Conjoint Experiment Robustness Checks.....	6
B.1	Intention to Comply as Outcome	6
B.2	Subset Analysis by Property Ownership	7
B.3	Subset Analysis by Residential Location	8
C	Lab-in-the-field Experiment Additional Analysis	9
C.1	Lab-in-the-field Experiment Subject Characteristics	9
C.2	Effect Estimates without First and Last Rounds.....	10
C.3	Effect Estimates without Seventh Round	11
C.4	Effect Estimates using Session Fixed Effects	12
C.5	Effect Estimates using Alternative Strategies to Control for Time Trends.....	13
C.6	Detailed Event History Estimates.....	14
C.7	Conjoint Analysis for Lab-in-the-field Subjects	15
C.8	Subset Analysis by Property Ownership.....	17
C.9	Power Analysis.....	18

A Conjoint Experiment Implementation Details

A.1 Conjoint Experiment Subject Characteristics

Table A-1: Summary Statistics for Conjoint Experiment Respondents

	Mean	Std. Dev.	of Respondents
Age (<i>Numeric</i>)	44.099	(11.948)	1786
Gender (<i>Binary, Male = 1</i>)	0.514	(0.500)	1786
Education (<i>Ordinal, 1-6</i>)	4.529	(0.891)	1786
CCP (<i>Binary</i>)	0.240	(0.427)	1786
Hukou (<i>Binary</i>)	0.859	(0.348)	1776
Income (<i>Ordinal, 1-23</i>)	6.132	(1.920)	1776
# of Apartments Owned (<i>Ordinal, 1-5</i>)	1.296	(0.645)	1776
Married (<i>Binary</i>)	0.896	(0.305)	1776

A.2 Balance Test for the Conjoint Experiment

Table A-2: Balance Test for SSI Conjoint Experiment Data

Dependent Variables: Model:	Age (1)	Gender (2)	Education (3)	CCP (4)	Hukou (5)	Income (6)	# Apts. Owned (7)	Married (8)
<i>Variables</i>								
Exemption	0.11 (0.41)	0.00 (0.02)	0.02 (0.03)	0.00 (0.01)	0.00 (0.01)	-0.06 (0.07)	-0.03 (0.02)	0.00 (0.01)
Penalty	0.23 (0.42)	-0.02 (0.02)	-0.02 (0.03)	0.01 (0.01)	0.01 (0.01)	0.10 (0.06)	-0.01 (0.02)	0.00 (0.01)
Citizen Input	0.85* (0.46)	0.03 (0.02)	-0.03 (0.03)	0.02 (0.02)	-0.01 (0.01)	-0.03 (0.07)	0.02 (0.02)	0.01 (0.01)
Top-down Supervision	-0.92** (0.44)	0.00 (0.02)	0.05 (0.03)	-0.02 (0.01)	-0.02 (0.01)	-0.02 (0.06)	0.00 (0.02)	0.00 (0.01)
Constant	43.94*** (0.74)	0.51*** (0.03)	4.52*** (0.05)	0.23*** (0.03)	0.87*** (0.02)	6.12*** (0.11)	1.30*** (0.04)	0.90*** (0.02)
<i>Fit statistics</i>								
Observations	3,432	3,432	3,432	3,432	3,408	3,408	3,408	3,408
R ²	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Clustered (Respondent) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A-3: Balance Test for Qualtrics Conjoint Experiment Data

Dependent Variables: Model:	Age (1)	Gender (2)	Education (3)	CCP (4)	Hukou (5)	Income (6)	# Apts. Owned (7)	Married (8)
<i>Variables</i>								
Exemption	-0.78** (0.31)	0.02 (0.01)	-0.01 (0.03)	-0.01 (0.01)	-0.01 (0.01)	0.11** (0.05)	0.01 (0.02)	0.00 (0.01)
Penalty	0.19 (0.41)	-0.02 (0.02)	-0.04 (0.03)	0.00 (0.02)	-0.01 (0.01)	0.00 (0.06)	0.03 (0.02)	-0.01 (0.01)
Citizen Input	-0.18 (0.38)	0.01 (0.02)	0.04 (0.03)	0.02 (0.01)	-0.01 (0.01)	0.03 (0.07)	-0.01 (0.02)	-0.01 (0.01)
Top-down Supervision	-0.07 (0.37)	0.00 (0.02)	-0.05 (0.03)	0.00 (0.01)	0.02 (0.01)	0.00 (0.06)	0.00 (0.02)	0.01 (0.01)
Constant	44.56*** (0.69)	0.51*** (0.03)	4.55*** (0.05)	0.25*** (0.02)	0.86*** (0.02)	6.08*** (0.12)	1.29*** (0.04)	0.89*** (0.02)
<i>Fit statistics</i>								
Observations	3,696	3,696	3,696	3,696	3,696	3,696	3,696	3,696
R ²	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Clustered (Respondent) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.3 Conjoint Experiment Screenshot

In this section, we show a real screenshot of a randomly generated paired comparison in the conjoint analysis and the Chinese translation of all values in the conjoint experiment.

There are no right or wrong answers, please just choose the policy that you personally prefer the most between the following two policies.

Round 1 out of 4:

	Policy 1	Policy 2
Is the first apartment exempted from the 0.5% tax rate?	Yes	Yes
How much interest is collected on the payment of late taxes when the apartment is sold?	3%	3%
Can citizens give input through conventional channels (e.g., public hearing, online commenting) about how revenue is spent?	No	Yes
Does the provincial auditing bureau audit property tax collection and punish city mayors for mismanagement of tax revenue?	No	Yes

Which policy would you choose?

Policy 1

Policy 2

How likely are you to comply with **Policy 1** on a scale of 1 to 5, 1 being never comply and 5 being always comply?

Never Comply

Rarely Comply

Sometimes Comply

Usually Comply

Always Comply

How likely are you to comply with **Policy 2** on a scale of 1 to 5, 1 being never comply and 5 being always comply?

Never Comply

Rarely Comply

Sometimes Comply

Usually Comply

Always Comply

请在以下两个房产税方案中选择您倾向的方案。不管您做出怎样的选择，这些选择都没有对错。

第一轮:

	方案 1	方案 2
市民是否可以通过常规性渠道（如听证会，网络意见收集）来建议房产税收入的使用方式？	有	没有
省级政府是否会每年定期监督市政府房产税收入的使用方式，并对违规的市长处以行政处分？	不会	会
是否每户家庭的第一套存量房免征0.5%的房产税	是	是
未按期交纳房产税在卖房时税款的滞纳金按应缴房产税的税款的年利率？	3%	15%

请选择：

方案1

方案2

如果实际政策是按照第一个房产税的方案，请问您是否会按时交纳房产税？在一个1至5的维度，1表明从不按时，5表明永远按时？

1 从不按时

2 很少按时

3 有时按时

4 经常按时

5 永远按时

如果实际政策是按照第二个房产税的方案，请问您是否会按时交纳房产税？在一个1至5的维度，1表明从不按时，5表明永远按时？

1 从不按时

2 很少按时

3 有时按时

4 经常按时

5 永远按时

Figure A-1: Conjoint Screenshot

B Conjoint Experiment Robustness Checks

B.1 Intention to Comply as Outcome

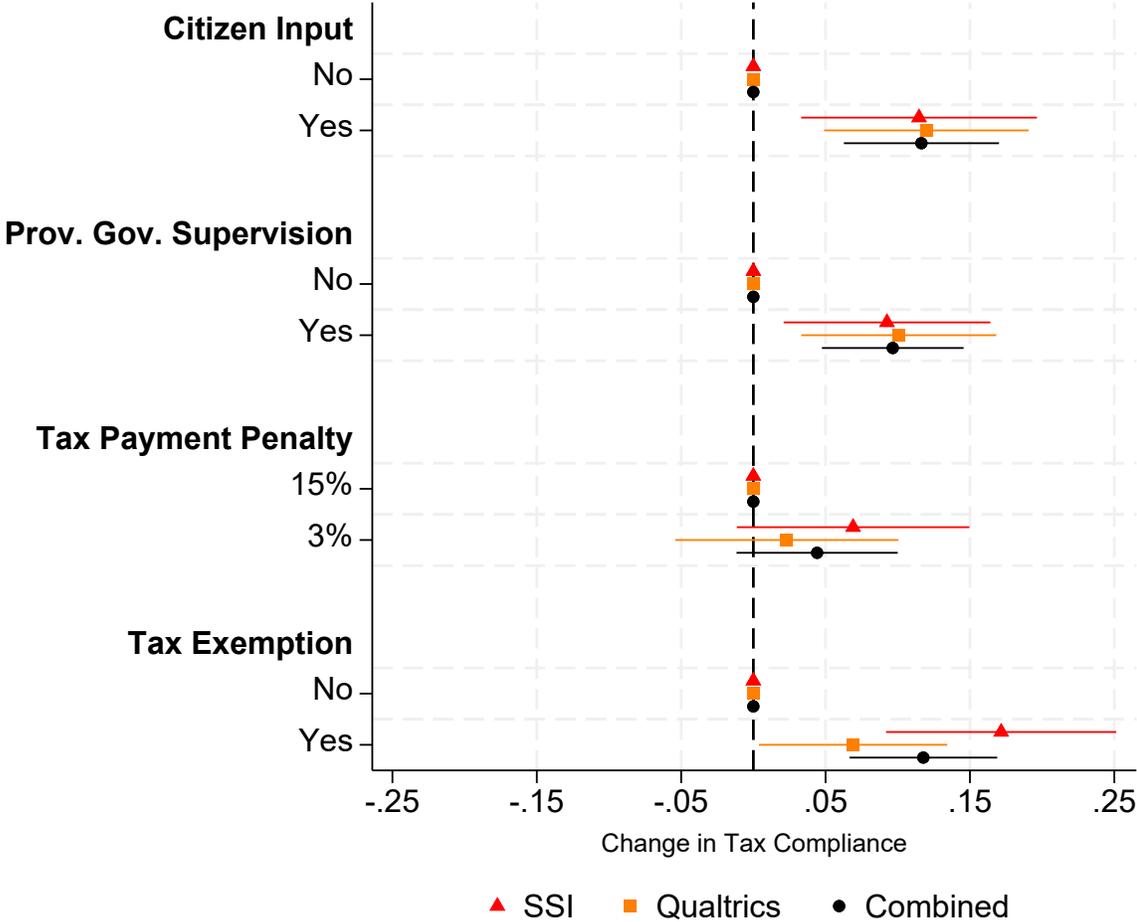


Figure B-2: Estimated AMCEs of conjoint attributes on intention to comply with tax policy (1-5 scale)

B.2 Subset Analysis by Property Ownership

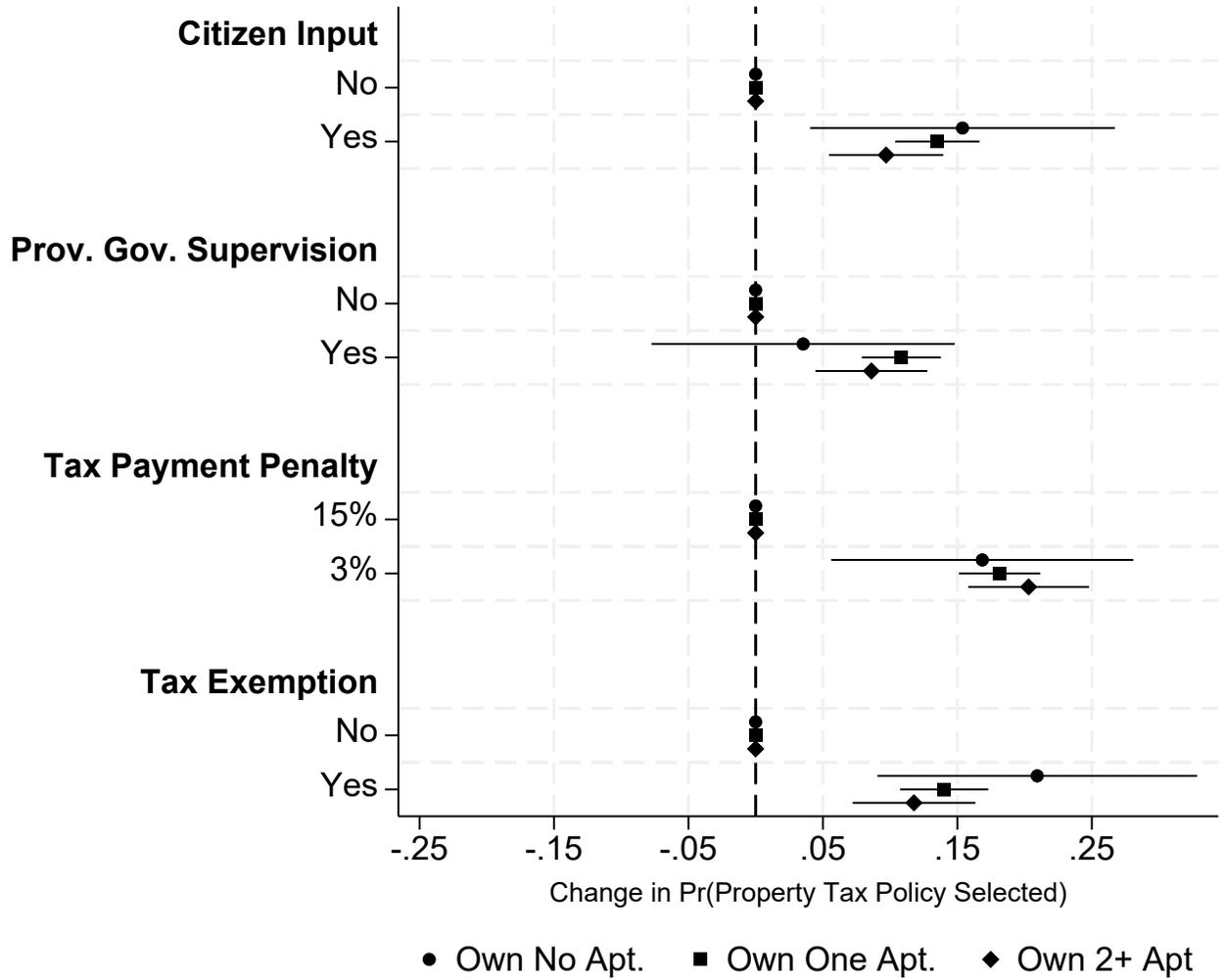


Figure B-3: Estimated AMCEs of conjoint attributes on tax policy preference, by apartment ownership

B.3 Subset Analysis by Residential Location

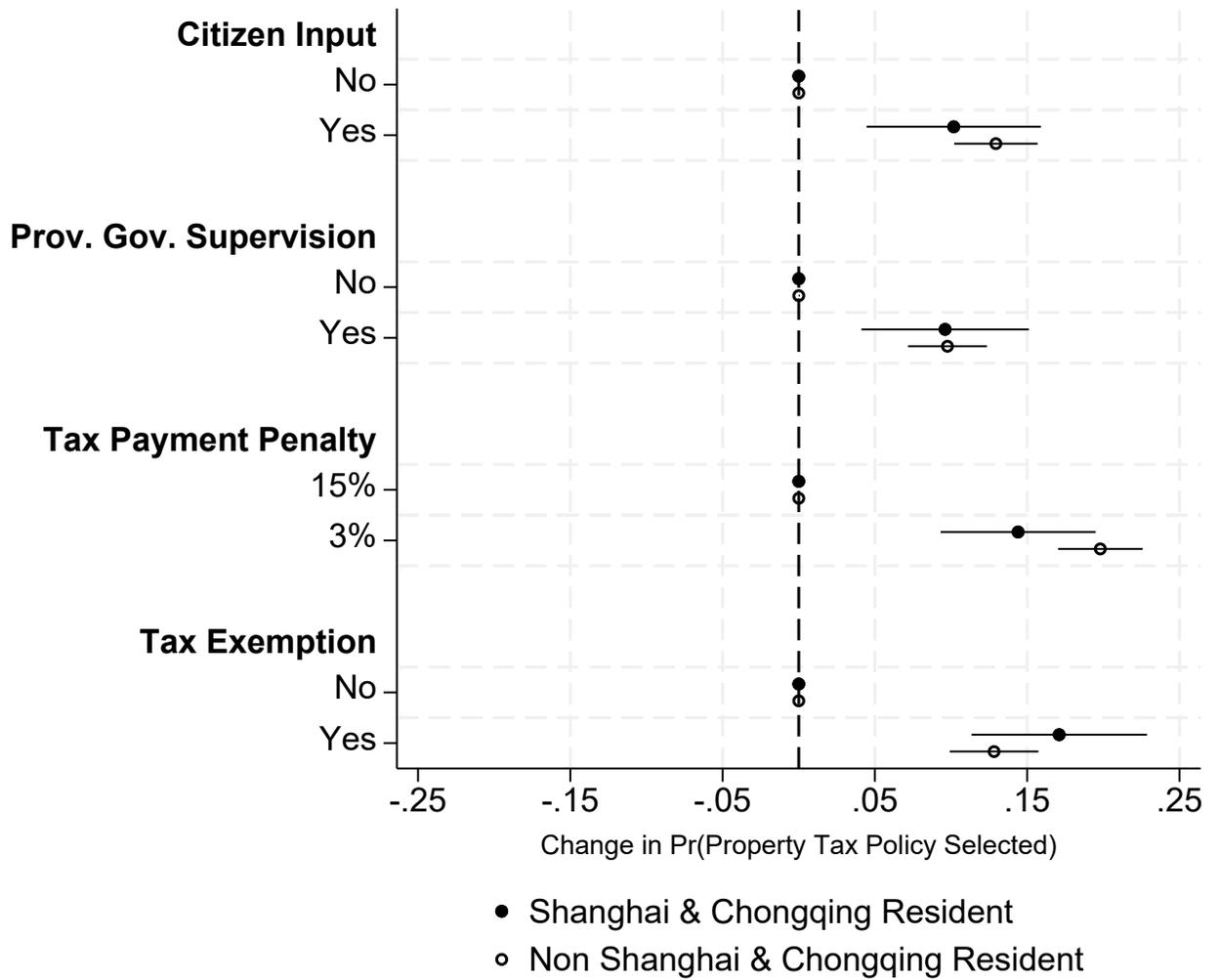


Figure B-4: Estimated AMCEs of conjoint attributes on tax policy preference, by residential location

C Lab-in-the-field Experiment Additional Analysis

C.1 Lab-in-the-field Experiment Subject Characteristics

Table C-4: Summary Statistics by Treatment Conditions

	Non-Responsive	Responsive	Non-Sanctioning	Sanctioning
Age	2.816	2.630	2.545	2.800
<i>(Ordinal, 1-5)</i>	(0.130)	(0.221)	(0.131)	(0.208)
Gender	0.579	0.407	0.485	0.320
<i>(Binary, Male = 1)</i>	(0.081)	(0.096)	(0.088)	(0.095)
Education	5.711	5.519	6.273	5.960
<i>(Ordinal, 1-6)</i>	(0.203)	(0.202)	(0.164)	(0.212)
CCP	0.500	0.370	0.455	0.680
<i>(Binary)</i>	(0.082)	(0.095)	(0.088)	(0.095)
Hukou	0.763	0.741	0.667	0.840
<i>(Binary)</i>	(0.070)	(0.086)	(0.083)	(0.075)
Income	18.395	18.593	19.531	18.160
<i>(Ordinal, 1-23)</i>	(0.756)	(0.787)	(0.801)	(1.029)
# of Apartments Owned	1.579	1.815	1.667	1.400
<i>(Ordinal, 1-5)</i>	(0.154)	(0.177)	(0.135)	(0.183)
Married	0.921	0.778	0.758	0.600
<i>(Binary)</i>	(0.044)	(0.082)	(0.076)	(0.100)
Obs.	38	27	33	25

C.2 Effect Estimates without First and Last Rounds

Table C-5: Impact of different types of government officials on tax compliance, excluding first and last rounds

Dependent Variable:	Tax Compliance		
	Average Effect		
Model:	(1)	(2)	(3)
<i>Variables</i>			
Non-Responsive	-0.058 (0.041)	-0.032 (0.044)	-0.044 (0.069)
Responsive	-0.044 (0.030)	-0.021 (0.045)	0.024 (0.057)
Non-Sanctioning	-0.030 (0.041)	-0.028 (0.058)	-0.089 (0.073)
Sanctioning	0.064 (0.059)	0.138** (0.060)	0.189** (0.076)
<i>Fixed-effects</i>			
Respondent	Yes	Yes	Yes
<i>Time trends</i>			
Common		Yes	
Individual			Yes
<i>Fit statistics</i>			
Observations	1,230	1,230	1,230
R ²	0.554	0.558	0.634

Clustered (Respondent) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Immediate effect estimates are not estimated because they do not make use of data from the first and last rounds.

C.3 Effect Estimates without Seventh Round

Table C-6: Impact of different types of government officials on tax compliance, excluding seventh round

Dependent Variable:	Tax Compliance		
	Average Effect		
Model:	(1)	(2)	(3)
<i>Variables</i>			
Non-Responsive	-0.096** (0.045)	-0.020 (0.057)	-0.024 (0.068)
Responsive	-0.026 (0.031)	0.021 (0.054)	0.019 (0.097)
Non-Sanctioning	-0.066 (0.055)	0.004 (0.065)	-0.036 (0.100)
Sanctioning	0.061 (0.059)	0.139* (0.072)	0.200 (0.129)
<i>Fixed-effects</i>			
Respondent	Yes	Yes	Yes
<i>Time trends</i>			
Common		Yes	
Individual			Yes
<i>Fit statistics</i>			
Observations	1,353	1,230	1,230
R ²	0.539	0.548	0.637

Clustered (Respondent) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Immediate effect estimates are not estimated because they do not make use of data from the 7th round.

C.4 Effect Estimates using Session Fixed Effects

Table C-7: Impact of different types of government officials on tax compliance, session fixed-effects

Dependent Variable:	Tax Compliance			
	Immediate Effect	Average Effect		
Model:	(1)	(2)	(3)	(4)
<i>Variables</i>				
Non-Responsive	-0.079 (0.069)	-0.083 (0.061)	-0.043 (0.048)	-0.039 (0.064)
Responsive	0.037 (0.082)	-0.019 (0.019)	-0.003 (0.037)	-0.005 (0.076)
Non-Sanctioning	-0.091* (0.049)	-0.045 (0.051)	-0.031 (0.054)	-0.087** (0.029)
Sanctioning	0.200* (0.094)	0.040 (0.028)	0.123*** (0.036)	0.192*** (0.051)
<i>Fixed-effects</i>				
Session	Yes	Yes	Yes	Yes
<i>Time trends</i>				
Common			Yes	
Session				Yes
<i>Fit statistics</i>				
Observations	246	1,476	1,353	1,353
R ²	0.161	0.102	0.109	0.119

Clustered (Session) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.5 Effect Estimates using Alternative Strategies to Control for Time Trends

Table C-8: Impact of different types of government officials on tax compliance, alternative time trends

Dependent Variable:		Tax Compliance						
		Average Effect						
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Variables</i>								
Non-Responsive	-0.043 (0.045)	-0.045 (0.046)	-0.068 (0.054)	-0.055 (0.057)	-0.039 (0.062)	-0.027 (0.063)	-0.087 (0.112)	-0.054 (0.128)
Responsive	-0.003 (0.043)	-0.006 (0.044)	-0.028 (0.053)	-0.015 (0.056)	-0.005 (0.060)	-0.036 (0.070)	0.023 (0.099)	0.013 (0.102)
Non-Sanctioning	-0.031 (0.059)	-0.033 (0.058)	-0.055 (0.066)	-0.042 (0.067)	-0.087 (0.067)	-0.092 (0.066)	-0.173** (0.086)	-0.174* (0.100)
Sanctioning	0.123** (0.057)	0.121** (0.058)	0.098 (0.065)	0.111* (0.066)	0.192** (0.081)	0.203** (0.091)	0.227** (0.102)	0.256** (0.116)
<i>Fixed-effects</i>								
Respondent	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Time trends</i>								
Common	Linear	Quadratic	Cubic	Quartic				
Respondent					Linear	Quadratic	Cubic	Quartic
<i>Fit statistics</i>								
Observations	1,353	1,353	1,353	1,353	1,353	1,353	1,353	1,353
R ²	0.552	0.552	0.552	0.553	0.634	0.676	0.715	0.745

Clustered (Respondent) standard-errors in parentheses

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

C.6 Detailed Event History Estimates

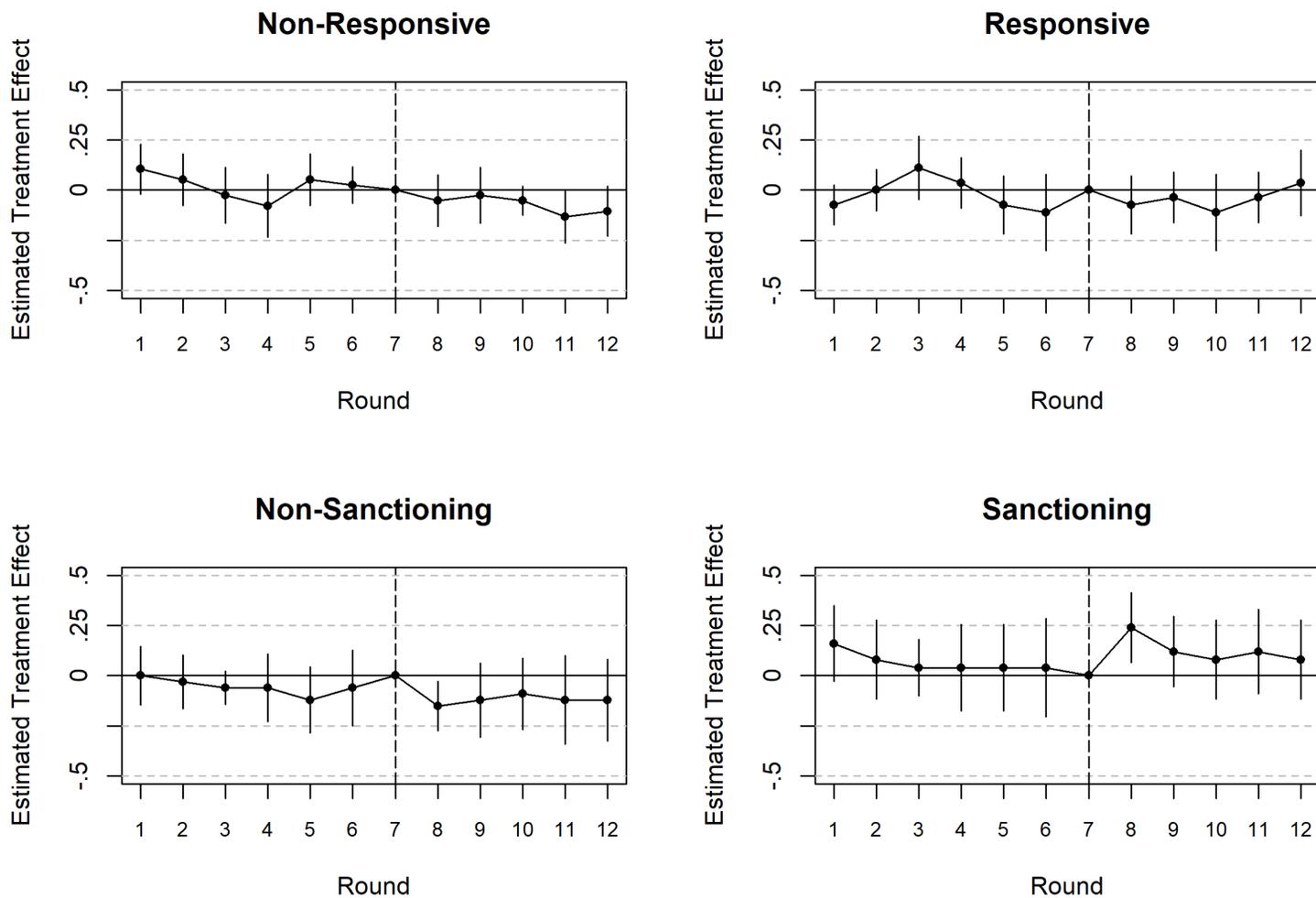


Figure C-5: Event history plot showing treatment effects estimated at every round for respondents in each treatment condition.

C.7 Conjoint Analysis for Lab-in-the-field Subjects

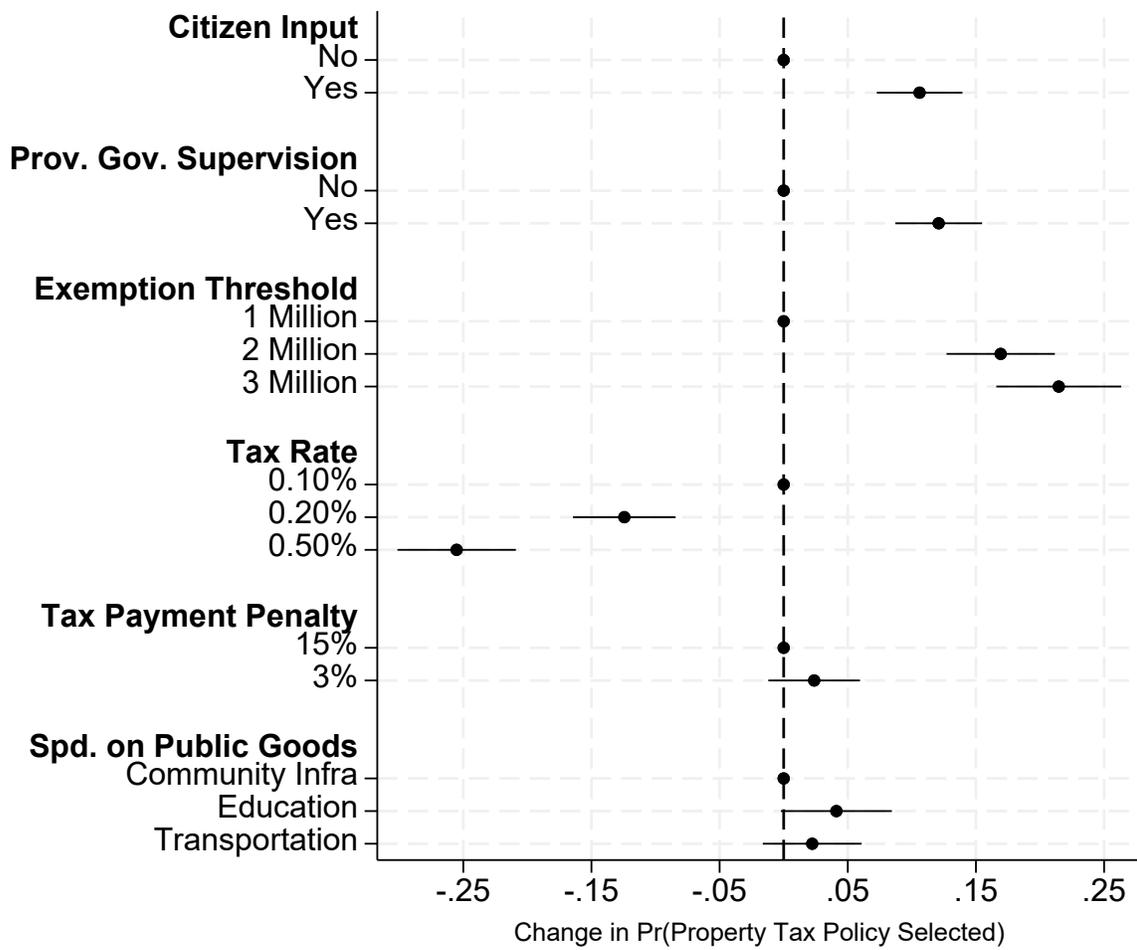


Figure C-6: Estimated AMCEs of conjoint attributes on tax policy preference, lab-in-the-field subjects

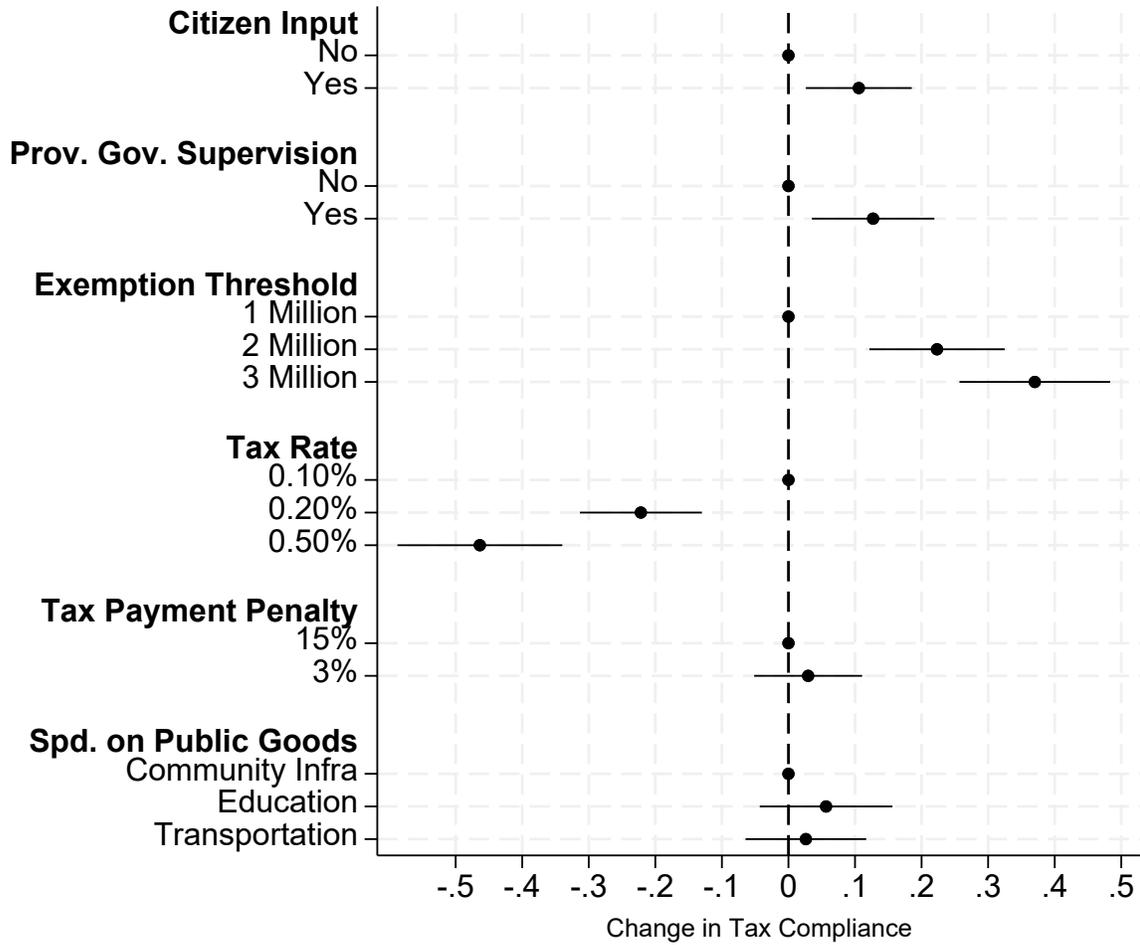


Figure C-7: Estimated AMCEs of conjoint attributes on intention to comply with tax policy (1-5 scale), lab-in-the-field subjects

C.8 Subset Analysis by Property Ownership

Table C-9: Impact of different types of government officials on tax compliance by property ownership and value

Dependent Variable:	Tax Compliance				
	Full Sample	# Apts. Owned		Property Value	
		1	2+	Bottom 75%	Top 25%
Model:	(1)	(2)	(3)	(4)	(5)
<i>Variables</i>					
Non-Responsive	-0.043 (0.045)	-0.047 (0.054)	-0.043 (0.079)	-0.039 (0.054)	-0.053 (0.076)
Responsive	-0.003 (0.043)	-0.003 (0.052)	0.004 (0.069)	-0.027 (0.058)	0.094 (0.061)
Non-Sanctioning	-0.031 (0.059)	0.012 (0.066)	-0.063 (0.099)	0.010 (0.072)	-0.153 (0.123)
Sanctioning	0.123** (0.057)	0.089 (0.064)	0.228 (0.152)	0.130* (0.068)	0.060 (0.127)
<i>Fixed-effects</i>					
Respondent	Yes	Yes	Yes	Yes	Yes
<i>Time trends</i>					
Common	Yes	Yes	Yes	Yes	Yes
Individual					
<i>Fit statistics</i>					
Observations	1,353	825	528	968	308
R ²	0.552	0.600	0.482	0.533	0.602

Clustered (Respondent) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.9 Power Analysis

Figures C-8 and C-9 present, for all five analytical approaches in Table 2, estimates of statistical power that our experiment could achieve given various sample sizes, levels of respondent heterogeneity, and magnitudes of the underlying treatment effect estimates. These estimates are obtained through a simulation exercise with the following parameters:

- Respondents are divided into a treatment and a control group, with the number of respondents per group N set to be various numbers between 20 and 100 to simulate different sample sizes
- The baseline control mean θ , i.e., the mean tax compliance rate when no intervention has taken place, is set to be 0.7.²
- Each respondent's compliance behaviors are drawn from Bernoulli distributions with individual means to simulate respondent heterogeneity. Each individual mean is drawn from a Beta distribution with mean $\mu_i = \theta$ and parameters α and β such that the resulting standard deviation is set to one of $\{0.2, 0.3, 0.4\}$
- Compliance behaviors are drawn for each respondent for each of 12 rounds.
 - For rounds 1-7 (control rounds), compliance behaviors are drawn from Bernoulli distributions with mean μ_i for all respondents
 - For rounds 8-12 (treatment rounds), compliance behaviors are drawn from Bernoulli distributions with means μ_i for control respondents and $[\mu_i + \tau]$ for treated respondents, with the treatment effect τ set to one of $\{0.2, 0.3, 0.4\}$ to simulate different treatment effect magnitudes.
- The simulation exercise is repeated 500 times. Each analytical model's statistical power is calculated as the proportion of these simulations in which this model returns a statistically significant result at the .05 level.

Our initial power analysis did not account for potential time trends; Figure C-8 presents estimated statistical powers from this initial analysis. Upon acknowledging a negative time trends from our empirical analysis, we reevaluate our experimental setup's power. Specifically, we adjust our Bernoulli means by $-0.005 \times \text{round number}$ for all respondent-rounds and re-run the simulation exercise.³ Figure C-9 presents estimated statistical powers from this modified analysis.

²The baseline control mean of 0.7 is based on an empirical estimate using our actual sample.

³The negative time trend of -0.005 is based on an empirical estimate using our actual sample.

In both Figures C-8 and C-9, the conventional threshold of sufficient power of 80% is denoted by horizontal dotted lines. We also denote $N = 30$, the approximate number of respondents in each of our actual sample's treatment groups, with a vertical dotted line. Note also that we estimate the standard deviation of our respondents' individual means to be around 0.36,

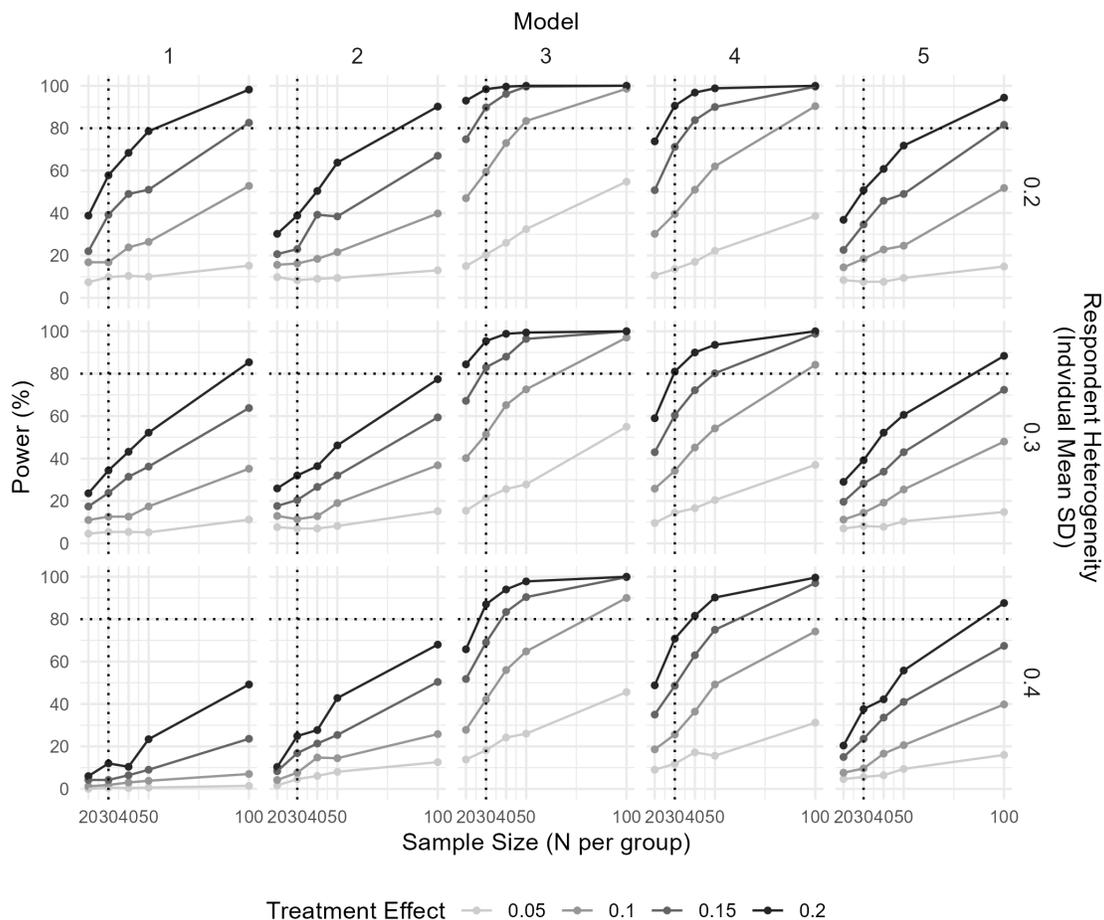


Figure C-8: Statistical power achieved by various models in Table 2 for various treatment effect magnitudes and different extent of respondent heterogeneity (measured as the standard deviation of the individual means of potential outcomes under control), *under no time trends*.

As can be seen in both Figures C-8 and C-8, analyses that focus on intermediate treatment effects, as in Models 1 and 2, are clearly underpowered. Even in the best-case scenario with a treatment effect as large as .2 (i.e., 20 percentage points), Model 1 only begins to deliver 80% statistical power at $N = 50$, as in the top-left cell of Figure C-8. Model 2 only begins to deliver the same level of power at $N = 100$. Given sample sizes, levels of respondent heterogeneity, and treatment effect magnitudes closer to what we found empirically, the two models achieve only about 20%-40% statistical power. This lack of power is unsurprising,

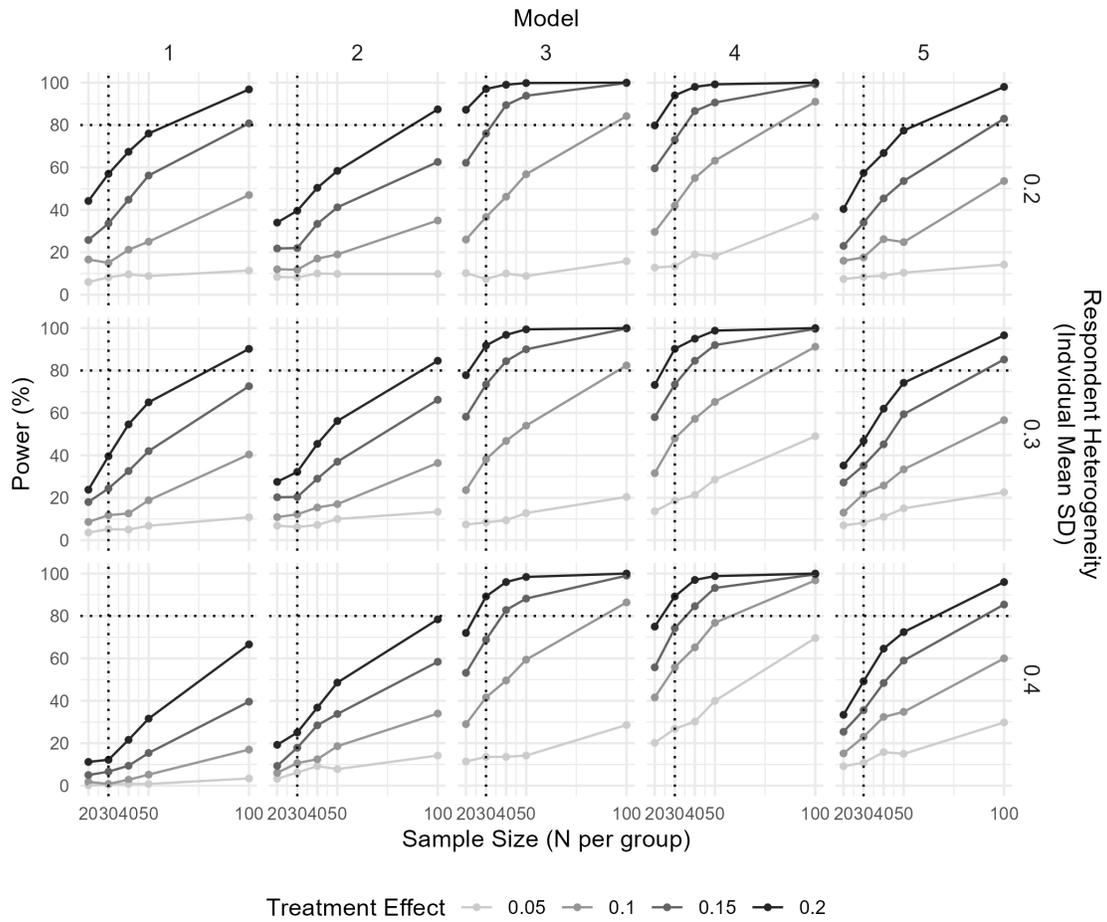


Figure C-9: Statistical power achieved by various models in Table 2 for various treatment effect magnitudes and different extent of respondent heterogeneity (measured as the standard deviation of the individual means of potential outcomes under control), *under a time trend of -.5 percentage points per round*.

given that these models use only two data points per respondent.

Our power analysis shows Model 3 to be the most statistically powered, with more than 90% power to detect treatment effects of magnitude .2, and between 70-90% to detect treatment effects of magnitude .15. However, as previously discussed, this model becomes biased when respondent compliance rates exhibit time trends.

Model 4, our preferred model given the existence of time trends, performs reasonably well for large treatment effects. Comparing between Figures C-8 and C-9, it is notable that its power actually increases when time trends exist. Under this scenario, it achieves more than 90% power to detect a large effect of .2 and above 70% power to detect a slightly smaller effect of .15 under most levels of respondent heterogeneity.

Finally, compared to Model 4, Model 5 achieves less statistical power, and in many cases even fails to outperform Models 1 and 2.

Ultimately, our power analysis indicates that our experimental data is best analyzed using Model 4, which takes advantage of repeated measures per respondent but also accounts for a time trend. It is sufficiently powered to detect fairly large treatment effects with magnitudes of .15 and above. With smaller treatment effects, however, even this model may run into power problems.