Understanding the Elasticity of Taxable Income: A Tale of Two Approaches*

Daixin He†, Langchuan Peng‡, and Xiaxin Wang (Job Market Paper)§

January 10, 2018

Abstract

This paper develops a framework to conduct the first formal comparison of two main approaches (the traditional tax reform approach and the recently developed bunching approach) to estimate the elasticity of taxable income (ETI), a central parameter in the public finance literature since Feldstein (1999). Using a novel panel of administrative personal income tax data from China and exploiting China's progressive monthly wage income tax schedule and a tax reform in 2011, we document two key differences between the ETI estimates using two approaches. First, the tax reform ETI estimates increase concavely over time, while the bunching ETI estimates are much more stable. Second, the tax reform ETI estimates (around 4 in the long-run) are much larger than the bunching ETI (around 0.5), and the difference is statistically significant. These stylized facts imply that very different behavioral responses are captured by the two approaches. To account for the stylized facts, we develop a simple model where individuals in each period have some probability to permanently change hours of work without paying other costs, but can temporarily adjust hours by paying additional costs. With stable wage rates, the two estimators should converge to the same underlying value. But with normal wage growth, the tax reform estimates converge to the true underlying parameter, whereas the bunching estimates can be far below the true figure. The findings imply that although the bunching approach have advantages in identification and application, the tax reform ETI estimates are generally more relevant for policy making due to the behavioral responses they are able to capture.

*Click here for the latest version. First draft: Oct 31, 2016. Xiaxin Wang is extremely grateful to Roger Gordon, Julie Cullen, and Eli Berman, for their invaluable guidance and consistent support. We are also grateful to Kate Antonovics, Prashant Bharadwaj, David Coyne, Gordon Dahl, Mitch Downey, Itzik Fadlon, Sieuwerd Gaastra, Alex Gelber, James Graham, Mark Jacobsen, Damon Jones, Claudio Labanca, Jeff Larrimore, Erzo F.P. Luttmer, Mark Machina, Karthik Muralidharan, Paul Niehaus, Mauricio Romero, Daniel Sacks, and audience in UCSD applied lunch seminar, UCSD summer job market seminar, All-California Labor Economics Conference (2016, UC Davis), WEAI (2017, San Diego), Young Economist Symposium (2017, Yale), NTA Annual Conference (2017), for helpful comments. The usual disclaimer applies.

†hedaixin@gmail.com; Chinese Academy of Social Sciences.

‡langchuan_peng@126.com; Institute of Economics and Finance, Nanjing Audit University.

§Corresponding author: xiaxinwang@ucsd.edu; Department of Economics, UC San Diego.

Keywords: elasticity of taxable income; bunching; tax reform.
JEL Codes: H20, H21, H24.

1 Introduction

The elasticity of taxable income (ETI) with respect to the marginal net-of-tax rate has been a central parameter in the public finance literature since Feldstein (1995, 1999). Feldstein (1999) shows that, even with tax evasion and avoidance, the ETI is a sufficient statistic for measuring the marginal efficiency cost of tax, and is therefore very useful for welfare analysis.[1] Empirically, there are two approaches to estimate the ETI, the traditional tax reform approach (e.g. Gruber and Saez (2002), Kleven and Schultz (2014)) and a recently developed bunching approach (e.g. Saez (2010), Chetty et al. (2011), Kleven and Waseem (2013)). While the traditional tax reform approach utilizes marginal tax rate changes induced by a tax reform to identify the behavioral responses summarized in taxable income changes, the bunching approach exploits the excess mass in the income distribution around a kink, where the marginal tax rate (MTR) changes, to identify behavioral responses local to a kink.

In addition, these two approaches have perceived strengths and weaknesses, respectively. The tax reform approach, as argued by Feldstein (1999), can generate an ETI estimate capturing behavioral responses at all margins (e.g. labor supply, tax avoidance, and tax evasion) to a tax change. Yet to apply this approach, it is necessary to have a tax reform. Moreover, how to address the endogeneity problem associated with this approach (caused by reversal causality and omitted variable bias) has been a central issue in literature. This approach generates estimates quite sensitive to instrumenting approach and regression specification (Saez et al. (2012)). By contrast, the bunching approach can be used in any setting with kinks or notches (where the average tax rate changes) in the tax code. The identification process can be transparently illustrated simply by showing the income distribution around a kink or a notch. Endogeneity is not a problem here, and the estimates are robust.[2] Due to these clear advantages, the bunching approach has recently been adopted in many settings (Kleven (2016)). But as an empirical strategy develped only recently, papers using the

[1] Chetty (2009) shows this is true if only a resource cost is involved in sheltering. He further shows that if sheltering is also associated with a transfer cost, then the elasticity of earnings and the resource cost of sheltering income from taxation are necessary to measure the deadweight loss of tax. Measuring the elasticity of earnings requires information on true earnings before sheltering; this requirement is rather demanding because even administrative tax data may not include incomes issued by cash and thus will underreport true earnings. Likewise, measuring the resource cost of sheltering also requires information unavailable in our administrative tax data. Due to these data limitations, we only focus on the ETI.

[2] Blomquist and Newey (2017) argue that, however, it may be necessary to know the functional form of the distribution of preference heterogeneity for the ETI to be identified using the bunching approach.

bunching approach have been more focused on "proof of concept" than policy evaluation. A deeper understanding of the behavioral responses captured by the bunching approach is clearly needed.

Although these two approaches are expected to measure the same underlying parameter, empirical evidence from previous papers suggests these two approaches could yield quite different estimates. Yet until now, there has been no formal comparison between these two approaches. Such an exploration is necessary, since otherwise we face a difficult choice between the two ETI estimates when trying to make policy implications out of them. In this paper, we compare the two approaches both empirically and conceptually, and show that they capture different behavioral responses to tax changes and thus have different policy implications.

Our empirical analysis focuses on China's personal wage income tax and is based on novel administrative data. The data cover administrative information on personal wage income tax, including monthly wage income of all wage/salary earners in a city with a population of 4-5 million, from June 2009 to December 2013. Desirable for our research purpose, China's personal wage income faces a graduated rate structure and had a reform on September 1st, 2011, after which a very wide range of income intervals have experienced marginal tax rate changes. This setting provides an ideal environment to apply both tax reform approach and bunching approach to estimate the ETI. One unique feature is that China imposes a *monthly* tax rate schedule on wage/salary, unlike most other countries where personal income faces a *yearly* tax rate schedule. This feature provides a rare opportunity to study the evolution of income responses to a tax change over a long enough time series, which is infeasible in many other settings.

To start with, we apply the standard tax reform approach (following Gruber and Saez (2002), Kleven and Schultz (2014)) and the standard bunching approach (following Saez (2010), Chetty et al. (2011), assuming no optimization frictions) to estimate the ETI. The standard tax reform approach renders an ETI of 2.42, robust to different instrumenting approaches. The bunching ETI estimates vary from 0.09 to 0.41 in several middle-high taxable income kinks. For lower taxable income kinks, there is no evidence of bunching, suggesting a zero ETI, as also observed in Kleven and Waseem (2013).[3] Consistent with previous studies that obtain ETI estimates using two approaches for the same country (US: Saez et al. (2012) and Saez (2010), Denmark: Kleven and Schultz (2014) and Chetty et al. (2011)), we find that in China, the tax reform ETI also seems systematically larger than the bunching ETI.

[3]For top kinks, since observations are too few to generate precise bunching estimates, they are not included in our bunching analysis.

The ETI estimates from the two standard approaches, however, are not yet directly comparable. The two standard approaches differ in two important aspects: time and scope. As for time, the standard tax reform approach renders 1-year, 2-year or 3-year ETI depending on specification, while the bunching approach renders an ETI with unclear time property, since it only requires cross-sectional or pooled cross-sectional data. As for scope, the tax reform approach generates a global ETI, while the bunching approach generates a local ETI for each kink. To make the two approaches more directly comparable, we develop a revised version for each approach to ensure that they obtain ETI estimates with the same property in time and scope.

Exploiting the advantage of our monthly income panel data and the tax reform in September 2011, we develop a revised tax reform approach that generates an ETI for each (three-month) period after the tax reform. The dynamic tax reform ETI estimates are consistent with the graphical evidence on the evolution of taxable income around the tax reform. Also relying on the monthly income panel data and the 2011 tax reform, we develop a revised bunching approach to explore the bunching responses to the introduction of post-reform kinks. Since there is no evidence of bunching at the post-reform kinks prior the tax reform, all bunching after the reform can be attributed to the reform. Therefore, we can simply apply the standard bunching approach to estimate an ETI for each post-reform period for each bunching kink. Then for each post-reform period, we derive a global bunching ETI estimate using observed ETI estimates at various bunching kinks. Here we adopt a revised version of the approach by Gelber et al. (2015) to estimate a common global ETI underlying all kinks. With optimization frictions, a common global ETI could generate different bunching behaviors (corresponding to the observed ETI estimates using standard bunching approach with no optimization frictions) at different kinks. We estimate the common underlying ETI and optimization frictions using the observed ETI estimates. If the assumption of a common underlying ETI with optimization frictions is reliable, then it should not only explain the observed bunching at middle-high kinks, but should also explain the lack of bunching at lower kinks. We use the estimated structural ETI and optimization frictions to confirm this. Finally, the revised bunching approach yields a sequence of dynamic global ETI estimates, which are compared to the dynamic tax reform ETI estimates.

We find two key differences between dynamic tax reform ETI estimates and bunching ETI estimates. First, the tax reform ETI estimates increase concavely over time, while the bunching ETI estimates are stable over time. Second, the tax reform ETI estimates (around 4 in the long-run) are much larger than the bunching ETI (around 0.5), and the difference is statistically significant. To account for the stylized facts, we develop a simple model where individuals in each period have some probability to permanently change hours of work without paying other costs, but can temporarily adjust hours by paying additional costs. The model implies that while the tax reform approach can capture the infrequent but permanent adjustment of hours of work to tax changes, the bunching approach generally reflect temporary adjustment. With stable wage rates, the two estimators should converge to the same underlying value. But with normal wage growth, the tax reform estimates converge to the true underlying parameter, whereas the bunching estimates can be far below the true figure.[4]

A welfare analysis based on our ETI estimates implies that the deadweight loss of China's current personal wage income tax is high and thus a further MTR decrease is desirable, as it would increase tax revenue and decrease deadweight loss. An evaluation of the 2011 tax reform reveals an interesting efficiency-neutral property, despite that the main objective of the reform is undoubtedly out of a redistribution concern, as it reduces the MTRs for lower earners and increases the MTRs for higher earners.

The major contribution of this paper is that it provides a first formal comparison of the two main approaches estimating the ETI.[5] Empirically, we document sharp contrasts between ETI estimates using two approaches around a tax reform. Conceptually and empirically, we show that the two approaches are measuring very different behavioral responses and thus are not interchangeable in general. The main findings of this paper imply that although the bunching approach have advantages in identification and application, the tax reform ETI estimates are generally more relevant for policy making due to the behavioral responses they are able to capture.

This paper is broadly related to papers reconciling different measures of the same policy relevant parameter. For example, Chetty et al. (2011) and Chetty (2012) try to reconcile micro and macro labor supply elasticities using adjustment costs and optimization frictions; Peterman (2016) tries to reconcile micro and macro estimates of the Frisch labor supply elasticity. Different from these papers, both the tax reform approach and the bunching approach yield *micro* estimates of a supposedly same parameter, and thus their sharp difference seems more puzzling.

In addition, this paper contributes to the large empirical ETI literature by providing the first Chinese evidence. The ETI estimates using both tax reform approach and bunching approach are both very large compared to those obtained in other countries (see Saez et al. (2012) and Saez et al. (2009) for a comprehensive review). There are several potential

[4] Although we emphasize the behavioral responses to tax changes via adjusting hours of work, other margins of change are possible, e.g. responses in income underreporting, intertemporal income shifting, and changes in labor participation.

[5] Recently we noticed that Miguel Almunia and Michael Best are working on a similar topic using UK data independently. We would appropriately cite their work once their draft is available.

reasons to account for the larger ETI estimates in China. First, China's personal income tax (PIT) system has a much cleaner tax base and a more salient tax schedule (the tax schedule does not depend on marital status, number of dependents, and is not inflation indexed), as opposed to the much more complicated PIT systems in countries like the U.S. and Denmark. As noted in previous literature, a simple tax code or tax reform would generate larger responses than a complicated one. Second, different personal income components (wage/salary, self-employment income, and other incomes) are taxed differently in China, as opposed to a universal personal income tax imposed in many other countries. This implies more space for income shifting between reported wage/salary and reported other incomes to save taxes. There could be other aspects (e.g. social culture, tax administration) underlying China's much larger behavioral responses to tax change.[6] Although this paper is not able to provide a comprehensive cross-country comparison, this could be a fruitful direction for future research.[7]

The remaining sections are organized as follows. Section 2 introduces China's personal income tax system, the 2011 tax reform, and the data. Section 3 applies the standard tax reform approach and the standard bunching approach to estimate the ETI. In section 4, after presenting a preliminary comparison between ETI estimates using two standard approaches in various countries, we develop a framework that compares the two approaches more directly. After documenting the differences between ETI estimates using two approaches, we explore the potential reason. Section 5 discusses the welfare implications from our ETI estimates and briefly evaluates the 2011 tax reform. Section 6 concludes.

[6] The differential definitions between taxable wage income and raw wage income would also partially account for the large elasticity in a mechanical way. Consider a person with a monthly wage income of 4,000 RMB. Suppose his wage income increases to 5,000 RMB in the next period, as a response to a decreasing tax rate. Then his raw income increases by 25%. Under the standard deduction 3,500 RMB, not considering other exemptions and deductions, his taxable income increases from 500 RMB to 1,500 RMB, which implies a 200% increase in the taxable income. Overall, the relatively large standard deduction to monthly income could account for a large ETI of the monthly wage income in China. Normally, the deductions and exemptions like those in the U.S. are not so large relative to income, especially when researchers focus on the high income earners, as many researchers do. Therefore, previous papers normally do not find very large difference between the ETI and the elasticity of raw income.

[7] Note that our estimates are obtained from only one city of China and thus should be cautiously interpreted on its representativeness when compared with estimates obtained in other countries.

2 China's personal income tax system and its 2011 reform

China's personal income tax system. China imposes a uniform nation-level personal income tax (PIT) schedule, with no additional PIT at the provincial or local level.[8] The PIT is levied on the individual rather than on the household level and is independent of the marital status and the number of dependents. Unlike the U.S., in China, there is no program like the Earned Income Tax Credit (EITC) for low income earners and the marginal PIT rate is always non-negative. The tax schedule is not indexed for inflation, which makes the bracket cutoffs more salient over time (due to its stable nominal value) than if it is inflation indexed.[9]

China's PIT deals with different income items separately (similar to Danish system, see Kleven and Schultz (2014)), unlike the US tax system which imposes a progressive rate oo the comprehensive taxable personal income. All income components can be divided into three types: (1) wage/salary income, (2) self-employment income, and (3) other incomes. According to the statutory schedule, wage/salary income is subject to a multiple-tier progressive rate structure, self-employment income is subject to a different multiple-tier progressive rate structure, and other incomes are subject to a proportional rate (in general 20%). In practice, however, self-employment income in general is not taxed following such progressive rate structure. Due to the absence of a reliable book-keeping, tax officials choose to enforce a pre-determined fixed amount self-employment income tax based on projected incomes for most self-employed businesses. The three types of incomes are also taxed on different time bases: wage/salary income is subject to a monthly schedule, self-employment income is subject to a yearly income schedule, and the other incomes are taxed each time the income is received. Another characteristic of China's PIT is that each income item is deducted separately instead of enjoying a deduction based on the comprehensive personal income.[10]

Since our main focus is the ETI w.r.t. the marginal tax rate and the self-employment income tax is not based on a rate structure, throughout this paper, we mainly focus on wage/salary income. Currently in China, for the majority of people, wage/salary income is their major income source. Bonuses are taxed differently from regular monthly wage by tax law, which could introduce complications both theoretically and empirically, as we discuss

[8] The personal income tax revenue (as well as the corporate income tax revenue) is shared between central (60%) and local (40%, in which normally 20% goes to province and 20% is retained locally) governments.

[9] In addition, there does not exist a comprehensive capital income tax in China, though many incomes that are counted as capital income in other countries are taxed under proportional tax rates (item 6, 7, 8, 9 in table A1).

[10] More details are discussed in Online Appendix A. Table A1 shows details on tax on all 11 personal income components.

in detail in Online Appendix B. However, our data show that too few incomes are taxed as bonuses to make bonuses an important concern and so we leave it out from our main analysis.[11]

Overall, China's personal income tax, in particular for the wage/salary income, is much simpler compared to many countries studied previously. Due to this, we expect to see much larger behavioral responses to tax changes in China. This is helpful for our empirical study of the behavioral responses to tax.

2011 PIT reform. During our data period (June 2009-December 2013), the 2011 PIT reform is the only major change in the PIT, which changed the standard deduction and the tax rate schedule for the wage/salary income and the self-employment income. There is no major change in other relevant taxes during this period.[12] The 2011 PIT reform proposal was passed on June 30 and was put into effect on September 1, 2011. In particular, for the wage/salary income, the monthly standard deduction increased from 2,000 RMB to 3,500 RMB, the 9-tier rate became 7-tier, and bracket cutoffs also changed. Figure 1 shows the personal income tax schedule for wages/salaries. It is clear that the 2011 PIT reform changed the marginal tax rate for a large scope of incomes. In particular, for taxable wage/salary incomes less than 4,500 RMB, marginal tax rates decreased; for those higher than 4,500 RMB, marginal tax rates increased whenever the marginal tax rates changed. These changes created substantial variations in the marginal tax rate faced by individuals and thus provided a good chance to examine behavioral responses.[13]

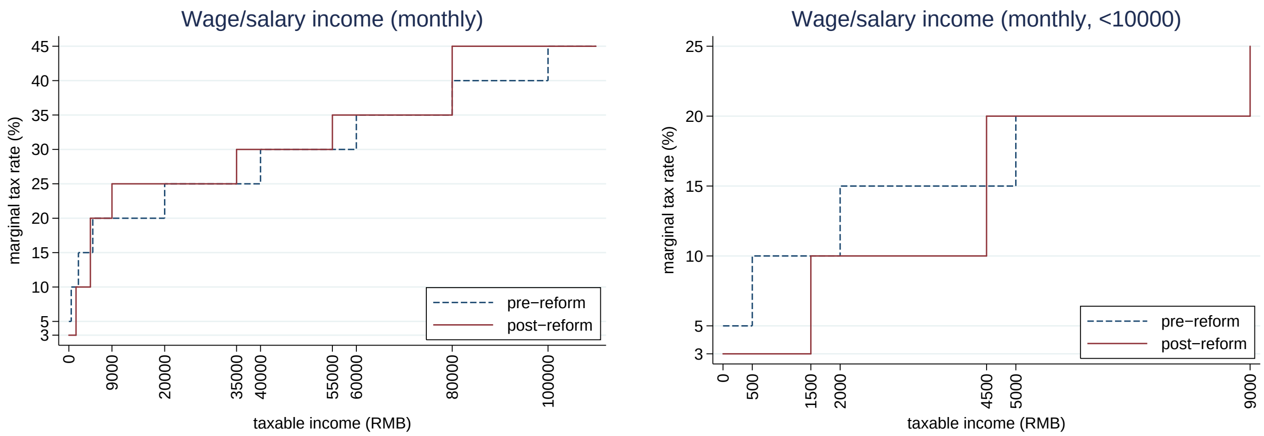
[11]Our data do not indicate which incomes are bonuses. Based on the actual tax rate and taxable income levels, we identify incomes following the "tax on bonuses" rule as bonuses. In 2013, only 0.46% (2,347 in all 505,159 individuals) of people have any bonuses in our data. But theoretically, people with annual income over 42,000 RMB should have part of their incomes issued as bonuses. In 2013, there are 192,893 individuals having annual earnings above 42,000 RMB. This fraction is very similar in other years. We are not entirely sure why there are so few people having bonuses. Perhaps many people receive bonuses in cash, as said in anecdotal evidence.

[12]Self-employed businesses do not need to pay corporate income tax (CIT), and the CIT rate is 25% for general firms and favorable rates apply for some specific firms. People need to pay social insurances (called *sijin* or *sanxianyijin* in China, including endowment insurance, medical insurance, unemployment insurance, employment injury insurance, maternity insurance, and housing fund, where maternity insurance is paid only by employers and the others are paid jointly by employees and employers). Even within a city, different firms may have different social insurance policies. There are occasional adjustments but no sharp change in social insurance policy in our sample city during our data period.

[13]For the self-employment income, the statutory tax schedule also changed (figure A1). But since most self-employment businesses pay a pre-determined fixed amount income tax, it is not clear how the statutory changes in marginal tax rate map into changes in the pre-determined fixed amount income tax.

Figure 1: Personal income tax schedule on wages/salaries

(a) (b)



Notes: The 2011 PIT reform proposal was passed on June 30 and put into effect on September 1st, 2011.

Personal income tax administrative data. Our personal income tax administrative data cover the whole population of a prefecture-level city in China from June 2009 to December 2013. The individual-level monthly panel dataset contains income and tax related information for all personal incomes subject to third-party reporting (mostly employer-reported). Variables include the unique individual ID, pre-tax monthly wage income, marginal tax rate, taxable income, tax liability, deductions and exempt incomes, sex, age, position, and occupation. No family-level information is available, as China's personal income tax does not depend on such information. Our sample city has a middle-sized population and a middle-high GDP level and so is not too unrepresentative of China.[14] The city has a population of 4-5 million and a 2014 GDP of 55-65 billion dollars (using 2014 exchange rate). Disposable income per capita of this city in 2014 falls in the range of 4,000-5,000 dollars. All wages/salaries data are included while the self-employment income data are unavailable to us.[15] The number of wage earners in each month varies from around 550,000 to 700,000. Table A2 shows that tax

[14]The city is not unrepresentative also in that it does not heavily rely on certain industries as compared to the national level. The fraction of its GDP coming from the three economic sectors are 6.9% for primary sector, 52.1% for secondary sector, and 41% for tertiary sector, as compared to 10%, 43.9% and 46.1% for the national level in 2013. And the fraction of employees hired by state-owned units is 20.1%, as compared to 16.6% for overall China. These statistics are calculated from China Statistical Yearbook and the statistical yearbook of our sample city.

[15]Since the wage/salary income is subject to third-party reporting, there should be minimal measurement error for this information. Importantly, employers report income for the employees even if their wages/salaries are below the standard deduction amount and do not need to pay any personal income tax. Self-employment income data are not reported to the department of local tax bureau that has all third-party reporting income and are thus not provided to us.

revenue components from various personal incomes in our sample city are comparable to the national figures. It is clear that the wage/salary income is the major personal income and in this paper we mainly focus on it. We restrict our sample to individuals between 18 and 60 to focus on the working age people.

3 Two standard approaches to estimate the ETI

3.1 Standard tax reform approach

The traditional approach to estimate the ETI exploits the tax rate changes induced by a tax reform. China's 2011 PIT reform created exogenous changes in the marginal tax rates for people in all income levels, thus providing enough variation to apply this approach. We follow the literature and apply the following first-difference specification to estimate the ETI e:

$$log\frac{z_{i,t+k}}{z_{it}} = e \cdot log(\frac{1-\tau_{i,t+k}}{1-\tau_{it}}) + \eta \cdot log(\frac{y_{i,t+k}}{y_{it}}) + f_t(z_{it_m}) + \Omega \cdot X_{it} + \alpha_t + \xi_{it},$$

where $log\frac{z_{i,t+k}}{z_{it}}$ is the growth rate of real taxable income (nominal taxable income adjusted by CPI) for individual i from time t to time $t+k$, τ_{it} is the marginal tax rate, y_{it} is virtual income defined below, η is income elasticity, X_{it} denotes dummies for demographic characteristics (age, sex, occupation, position), α_t are month fixed-effects. $\xi_{it} = \varepsilon_{i,t+k} - \varepsilon_{it}$, where ε_{it} is the error term of the function determining $log\, z_{it}$. We follow the common practice in literature to define taxable income z_{it} in a way that the tax base is constant throughout the period.[16] Without this adjustment, the dependent variable changes mechanically as the definition of the tax base changes; with this adjustment, what we estimate is entirely due to the MTR change. While most previous papers use yearly data, we use monthly data, since a monthly tax schedule is applied to wage/salary income in China, and we adjust the specification accordingly. Since our data only cover two years before and two years after the reform, our preferred regression uses 12-month (1-year) difference. 12 months is an appropriate choice since it is long enough to allow wage adjustment and not too long given our data covering period.[17]

[16] Taxable income is defined as raw income - standard deduction - other deductions - tax-exempt incomes. We apply the post-reform tax base by assuming the pre-reform observations are subject to post-reform standard deduction, as the only change in the tax base during the 2011 tax reform is change of the standard deduction.

[17] Our data have similar structure to Ito (2014), who uses monthly electricity consumption data and estimates the effects of marginal price and average price on the monthly electricity consumption. So we follow his specification in many aspects, i.e. using 12 month first-difference specification, using middle-time taxable income to construct instruments, using decile-by-month fixed effects to control for heterogeneous underlying

Since tax rates are a function of taxable income, the log change in the net-of-tax rate is clearly endogenous. To address this problem, $log(\frac{1-\tau_{i,t+k}}{1-\tau_{it}})$ is instrumented using $log(1-\tau_{i,t+k}(\widetilde{z_{it}}))-log(1-\tau_{it}(\widetilde{z_{it}}))$. This instrument computes the predicted net-of-tax rate change at a taxable income level $\widetilde{z_{it}}$. The idea of such an instrument strategy is to just use exogenous changes in tax laws to provide identification. The traditional practice in literature (e.g. Gruber and Saez (2002), Kopczuk (2005)) is to use $\widetilde{z_{it}} = z_{it}$. However, as widely recognized, z_{it} is likely to be correlated with ξ_{it} because the mean reversion of income creates a negative correlation between ε_{it} and $\xi_{it} = \varepsilon_{i,t+k} - \varepsilon_{it}$. Some strategies are thus proposed to address this problem.

Our preferred instrument strategy follows that used in Ito (2014) and Blomquist and Selin (2010) to use the taxable income in the middle time between t and $t+k$ to generate the simulated MTR change. In our case, $k = 12$ and so the middle time is $t_m = t + 6$ and the instrument is based on $z_{i,t+6}$. As shown in Ito (2014) and Blomquist and Selin (2010), this instrument is not systematically affected by the mean reversion problem because $\varepsilon_{i,t+12}$ and ε_{it} do not directly affect $z_{i,t+6}$. If there is no serial correlation, $\varepsilon_{i,t+6}$ and $\xi_{it} = \varepsilon_{i,t+12} - \varepsilon_{it}$ are clearly uncorrelated. When there is serial correlation, Blomquist and Selin (2010) show that $cov(\varepsilon_{i,t+12} - \varepsilon_{it}, \varepsilon_{i,t+6}) = 0$ as long as the serial correlation depends only on the time difference between the error terms. The intuition is that since $\varepsilon_{i,t+6}$ is equally spaced from $\varepsilon_{i,t+12}$ and ε_{it}, it would be correlated with them in the same manner. Alternatively, Weber (2014) proposes an instrument approach to mitigate the mean reversion problem. She argues that using lagged terms of z_{it} instead of z_{it} itself to construct the predicted MTR would render instruments that are strictly more orthogonal to the error term than traditional instrument. We apply this approach to use one-year lagged z_{it} to construct the alternative instrument.[18] [19]

Most ETI literature simply ignores the income effect since previous literature generally

changes in taxable income growth for different income levels.

[18]We use the middle time taxable income based instrument rather than the Weber-type instrument as our preferred instrument strategy for three reasons. First, the Weber-style instrument does not guarantee a strictly exogenous instrument while the middle time taxable income based instrument used in Ito (2014) and Blomquist and Selin (2010) does under reasonable assumptions. Second, the Weber-style instrument strategy greatly shrinks our sample period that can be used for regression, while the middle time taxable income based instrument does not. Third, the Weber-style instrument faces a trade-off between two requirements that make an instrument valid. That is, a longer lag of taxable income based instrument will make the exclusion restriction more reliable (since serial correlation of error terms will be weaker) and the weak IV problem more acute (since the first stage result will be weaker). But no criterion is proposed on how to decide between them. Weber simply assumes a longer lag to be more orthogonal, given its first-stage result is not weak. By contrast, the middle time taxable income based instrument could satisfy the exclusion restriction assumption under reasonable assumption, without sacrificing the first-stage. In subsequent regressions, we only focus on middle time taxable income based instrument.

[19]We are unable to use the two-year lag of z_{it} to construct the instrument in our data since this would leave too few pre-reform months (only three).

finds it small (e.g. Gruber and Saez (2002) for the US and Kleven and Schultz (2014) for Denmark). But different economies could have different sizes of income effects. And as noted by Gruber and Saez (2002), it is theoretically unclear what sign to expect for the income effect estimates for constructs such as broad or taxable income. So we explicitly examine the size of income effects in China. Our empirical estimates start with a specification without income effects, and then control for the log difference of virtual income, where virtual income $y_{it} \equiv \tau_{it} \cdot z_{it} - T_t(z_{it})$, with $T_t(.)$ denoting tax liability, following Kleven and Schultz (2014), Blomquist and Selin (2010), Bastani and Selin (2014), and Jantti et al. (2015).[20] The instrument for virtual income is constructed in a similar spirit to that used for $log(\frac{1-\tau_{i,t+k}}{1-\tau_{it}})$. When income effect is important, the estimate e is an uncompensated elasticity due to budget set linearization implied by the virtual income formulation.[21] If income effect is small and unimportant, we can use the specification with no income effects and interpret e as the compensated elasticity.

There are many ways to define $f_t(z_{it_m})$. For example, we can include flexible polynomial functions of z_{it_m}. But to avoid imposing a functional form assumption, we take a non-parametric approach. In particular, we include a set of decile dummies of taxable income for each t_m. By doing so, we have a set of decile-by-month fixed effects.[22] Such flexible controls of z_{it_m} account for heterogeneous income growth rates of different income levels. When we use the Weber-type instrument, we accordingly include a set of decile-by-month fixed effects based on $z_{i,t-12}$.

Regressions are weighted by middle-time taxable income ($z_{i,t+6}$) or lagged taxable income ($z_{i,t-12}$) depending on the instrument used.

Table 1 shows the regression results. Columns 1 and 2 are our preferred results, using middle-time taxable income based instruments. Column 1 shows the estimate without income effects.[23] The point estimate of ETI is 2.423 and is statistically significant at 1% level. Column 2 includes the income effect and shows that it is small and statistically insignificant in

[20]As noted by Kleven and Schultz (2014), modeling the income effect in terms of virtual income deviates from some previous taxable income studies (e.g. Gruber and Saez (2002)), where the income effect is specified simply in terms of after-tax income $z_{it} - T_t(z_{it})$. But as noted by Blomquist and Selin (2010), Bastani and Selin (2014), Jantti et al. (2015), the virtual income specification more closely follows the labor literature of specifying income effects and therefore is widely adopted in these recent taxable income studies.

[21]The compensated elasticity is then $\zeta^c = e - \eta\frac{(1-\tau)z}{y}$, where y is virtual income and η is elasticity w.r.t. virtual income (Blomquist and Selin (2010)). In Gruber and Saez (2002), they use after-tax income as a proxy of virtual income, i.e. $y = (1-\tau)z$, and they have $\zeta^c = e - \eta$.

[22]Using percentile-by-month fixed effects renders very similar results.

[23]First-stage results are strong in all columns.

Table 1: Estimates of ETI for wage/salary income using tax reform approach

Dep. Var.: Δlog(taxable income)	[1]	[2]	[3]	[4]
Δln(1-τ)	2.423***	2.534***	2.739***	4.023**
	(0.188)	(0.645)	(0.500)	(1.646)
Δln(virtual income)		0.117		1.705
		(0.712)		(2.178)
Instrument based on	taxable income in t+6		taxable income in t-12	
Observations	1,210,376	1,210,376	572,509	572,509
First-stage F-stat for Δln(1-τ)	3647	1808	995	497.9
First-stage F-stat for Δln(virtual income)		161.8		43.44

Notes: The table shows elasticity estimates based on 2SLS regressions, where standard errors (shown in parentheses) are clustered by individual. *** p<0.01, ** p<0.05, * p<0.1. All specifications include dummies of age, gender, occupation, position, middle-month taxable income (or one year lag taxable income) decile-by-month fixed effects, and base-month fixed effects. Regressions are weighted by middle-month taxable income (or one year lag taxable income).

our case, consistent with most previous literature. Columns 3 and 4 show results using Weber-type instruments. Column 3 shows a compensated ETI estimate similar as our preferred specification. Column 4 shows a large yet statistically insignificant income effect. Therefore, we regard 2.423 as our compensated ETI estimate using the standard tax reform approach and ignore income effects throughout our paper. These estimates are in line with the ETI estimates obtained in our revised tax reform approach below.

3.2 Standard bunching approach

Due to the increasing availability of administrative tax returns data, there has been a surge of research using the bunching approach to estimate compensated elasticities. Notably, Saez (2010) and Chetty et al. (2011) use bunching at kinks and Kleven and Waseem (2013) use bunching at notches to uncover compensated elasticities and the underlying structural elasticities. In this section, we use the standard bunching approach developed by Chetty et al. (2011) using pooled cross-sectional data to estimate the ETI e without considering optimization frictions.

3.2.1 Conceptual framework and empirical strategy

The standard bunching approach to estimate the elasticity of taxable income can be briefly described as follows.[24] Consider individuals with preferences defined on after-tax income and before-tax income. The utility function is $u(z-T(z), \frac{z}{n})$, where z is earnings, $T(z)$ is tax over earnings, and n denotes ability. Suppose the initial marginal tax rate is τ_1 and an increase in the marginal tax rate starting at taxable income K is introduced, bringing marginal tax rate to $\tau_2 = \tau_1 + \triangle\tau$ for taxable income above K. Under this two-tier tax schedule, all individuals originally choosing K or less are not affected. The individual whose indifference curve was tangent to the original budget line at $K + \triangle z$ now has indifference curve tangent to the upper part of the two-tier budget line at K. This individual is called the marginal buncher because all the individuals initially locating between K and $K + \triangle z$ now would choose K. All the individuals initially choosing $(K,\ K + \triangle z)$ are called bunchers. Some individuals that originally chose more than $K + \triangle z$ may now choose taxable income between K and $K+\triangle z$.[25] Thus, in theory a convex kink at K would generate excess bunching at K. Assume e, the elasticity of taxable income with respect to net-of-tax rate, is constant for individuals around the kink K, then by definition we have

$$e = \frac{\triangle z/K}{\triangle\tau/(1-\tau_1)},$$

where only $\triangle z$ needs to be identified to estimate e. Denote the excess bunching amount by B, we have

$$B = \int_K^{K+\triangle z} h_0(z)dz = h_0(\bar{z}) \triangle z \simeq h_0(K) \triangle z,$$

where $h_0(z)$ is the density function of taxable income when there is a constant marginal tax rate τ_1 throughout the distribution. The second equality is due to the mean value theorem for integrals, and $\bar{z} \in [K,\ K+\triangle z]$. When $\triangle z$ is small, $h_0(\bar{z})$ is approximated using $h_0(K)$. In theory, $h_0(K)$ is the density function at point K, while empirically we estimate the density on bins with width W. So we modify the above relation as

$$B \simeq \frac{h_0^W(K) \triangle z}{W},$$

where $h_0^W(K)$ is the density associated with bins of width W. Plugging it back to the

[24]The standard bunching approach tends to ignore the income effect and we follow this tradition in this paper. First, since we find small income effect in our case as in most previous studies, this seems reasonable. Second, Bastani and Selin (2014) use numerical simulation to show that, even when the kink is very large (and the income effect is thus plausibly large), income effects are unlikely to bias the ETI estimates from the standard bunching approach.

[25]See Kleven (2016) or Saez (2010) for a graphical illustration.

definition of elasticity, we have

$$e \simeq \frac{B/h_0^W(K)}{\frac{K}{W}\cdot\frac{\triangle\tau}{1-\tau_1}}.$$

Then it suffices to estimate $b \equiv \frac{B}{h_0^W(K)}$, the fraction of excess bunchers normalized by the counterfactual density.

To estimate b, we apply the standard approach used in Chetty et al. (2011). Observations around kinks are first grouped into bins with width W. Denoting by c_j the number of observations and z_j the taxable income relative to kink K in bin j, we fit a flexible polynomial of order q to the bin counts in the empirical distribution[26], omitting the excluded region $(z_L,\, z_U)$,[27] by estimating regression:

$$c_j = \sum_{i=0}^{q} \beta_i^0 \cdot (z_j)^i + \sum_{i\in(z_L,\, z_U)} \gamma_i^0 \cdot \mathbf{1}[z_j = i] + \varepsilon_j,$$

where γ_i^0 is a bin fixed effect for each bin in the excluded region. The initial estimate of the counterfactual distribution is the predicted values from the above regression by setting all the dummies in the excluded region to zero: $\widehat{c_j}^0 = \sum_{i=0}^{q} \widehat{\beta_i^0} \cdot (z_j)^i$. The initial estimate of excess bunching, defined as the difference between the observed and counterfactual counts within the excluded region, is $\widehat{B}^0 = \sum_{j\in(z_L,\, z_U)}(c_j - \widehat{c_j}^0)$. $\widehat{B}^0$ might overestimate $\widehat{B}$ because it does not account for the fact that the additional individuals at the kink come from points to the right of the kink. Hence the estimated counterfactual is likely to be based on an underestimate of individuals that would have been observed without the kink. Following Chetty et al. (2011), we address this concern by shifting the counterfactual distribution to the right of the kink upward until it satisfies the constraint that the number of observations in the counterfactual distribution is equal to the number of observations in the observed distribution. In particular, the final estimate of the counterfactual distribution is the predicted values $\widehat{c_j} = \sum_{i=0}^{q} \widehat{\beta_i} \cdot (z_j)^i$ from the following regression:

$$c_j \cdot \left(1 + \mathbf{1}[j \geq z_U] \cdot \frac{\widehat{B}^0}{\sum_{j=z_U}^{\infty} c_j}\right) = \sum_{i=0}^{q} \beta_i \cdot (z_j)^i + \sum_{i\in(z_L,\, z_U)} \gamma_i \cdot \mathbf{1}[z_j = i] + \varepsilon_j. \tag{1}$$

[26] In practice, we take the seventh-degree polynomial, following Chetty et al. (2011).

[27] The excluded region is the region around the kink where excess bunching happens. In the case of kinks, the excluded region is typically determined visually, while in the case of notches, there is additional moment to help determine the bounds of the excluded region. See Kleven (2016) for a comprehensive review.

Then we obtain $\widehat{B} = \sum_{j \in (z_L, z_U)} (c_j - \widehat{c_j})$. The empirical estimate of b is given by

$$\widehat{b} = \frac{\widehat{B}}{(\sum_{j \in (z_L, z_U)} \widehat{c_j})/N}, \tag{2}$$

where N is the number of bins in the excluded region. Following Chetty et al. (2011), the standard error for $\widehat{b}$ is bootstrapped. We randomly draw from the estimated vector of errors $\widehat{\varepsilon_j}$ in (1) with replacement and generate a new set of counts and apply the above technique to calculate a new set of estimates $\widehat{b}^k$s. Define the standard error of $\widehat{b}$ as the standard deviation of the distribution of $\widehat{b}^k$s. Finally, e can be obtained as $\widehat{e} \simeq \frac{\widehat{b}}{\frac{K}{W} \cdot \frac{\triangle \tau}{1-\tau_1}}$, with standard error computed using the delta method $std(\widehat{e}) \simeq \frac{std(\widehat{b})}{|\frac{K}{W} \cdot \frac{\triangle \tau}{1-\tau_1}|}$.

3.2.2 Standard bunching estimates

Unlike previous bunching papers that use a full sample, our main analysis applies the bunching method to a decimal sample (i.e. dropping observations with taxable income exactly at a round number and keeping only those with decimal values). This restriction is made in order to address the irregular bunching at non-kink places observed in our data. Previous literature accounts for regular bunching patterns at non-kink numbers by adding indicators of different "rounder" numbers.[28] Yet this approach cannot address the irregular bunching patterns in our case. Restricting sample to decimal TI values well addresses this issue and reveals reliable bunching patterns at kinks and exclude any bunching at non-kink places. In Online Appendix C, we show the bunching patterns using full sample and decimal sample and discuss the sample restriction in detail.[29]

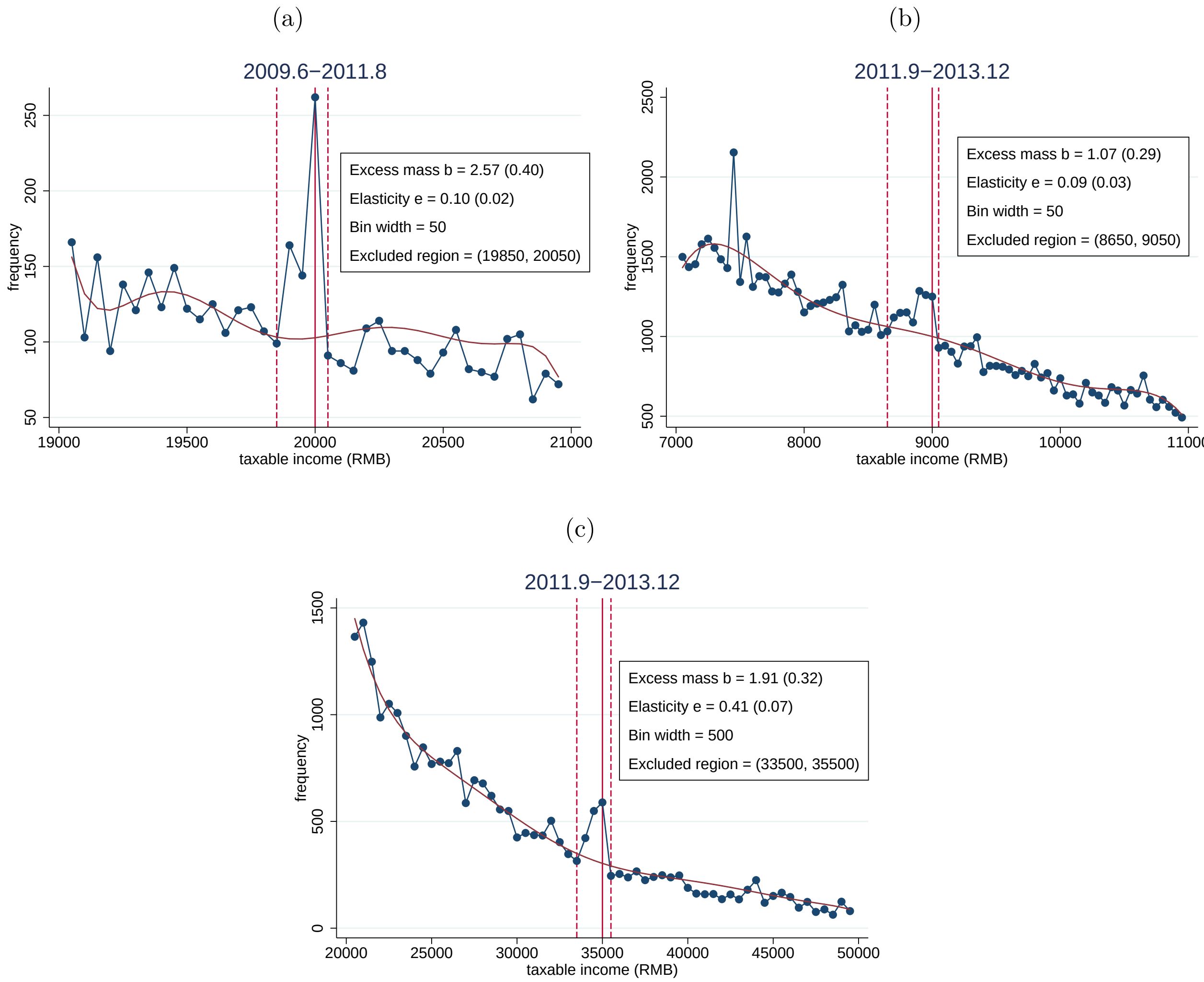
Given the large sample size of our dataset, restricting to a decimal sample would still render precisely estimated ETI for each kink. Figure A4 shows clear bunching at pre-reform kink 20,000 RMB before the tax reform and at post-reform kinks at 9,000 RMB and 35,000 RMB after the reform. Like Kleven and Waseem (2013), we find no evidence of bunching at bottom kinks, suggesting a zero ETI there, and observations are too few to generate precise

[28] Ignoring this rounder-number bunching behavior could have the standard bunching approach overestimate the ETI. Kleven and Waseem (2013) under the notch setting propose a way to address the rounder-number bunching problem by including an indicator for rounder numbers (i.e. multiples of 5K, 10K, 25K, and 50K) when estimating the counterfactual density function. Devereux et al. (2014) follow such approach to estimate the ETI of corporate income tax in the UK and Best and Kleven (2016) adopt a similar approach to deal with the rounder-number bunching for house prices. Some bunching analyses simply ignore the rounder-number bunching problem, probably because in their specific cases the rounder-number bunching problem is not salient (e.g. Chetty et al. (2011) and Saez (2010)).

[29] Admittedly, there may be concern that the decimal sample could underestimate the ETI since it might exclude taxable incomes adjusted to exactly at the kink more than those with other integer values. We address this concern by showing that using full sample would generate similar dynamic pattern of bunching estimates in Online Appendix F.

bunching estimates at top kinks. Thus, we focus on the middle-high TI kinks to apply the standard bunching approach.

Figure 2: Standard bunching estimates of ETI



Notes: The solid smooth curve depicts the estimated counterfactual distribution omitting the obervations in the excluded region, as specified by the area between the dashed lines.

Estimates. Figure 2 shows the estimates of the excess mass and the elasticity of taxable wage income. The solid dotted line depicts the observed distribution and the solid smooth line shows the estimated counterfactual distribution omitting the observations in the excluded region. For kinks at 20,000 RMB and 9,000 RMB, width of bins is 50 RMB, while for kink at 35,000 RMB, where observations are much scarce, width of bins is 500 RMB. The observed elasticity for pre-reform kink 20,000 RMB is 0.10. The observed elasticities for post-reform kinks 9,000 RMB and 35,000 RMB are 0.09 and 0.41, respectively. The elasticities are all statistically significant at 1% level.[30] As a placebo test, in Online Appendix D, we show that there was no bunching at all before a new kink was imposed, and that bunching disappeared within a short time after an old kink was abolished.

3.2.3 Who are the bunchers?

In this section, we provide the first formal test of the key assumption of the bunching approach, and then examine personal characteristics (i.e. sex, age, occupation, position) of bunchers versus non-bunchers.

Testing the key assumption of bunching approach. The key assumption of the bunching approach is that excess bunchers mainly come from those that could have earned slightly more than the income associated with the kink. This assumption determines whether the bunching approach measures the income adjustment behavior as it claims but has never been formally tested. An alternative possibility is that a non-negligible portion of the excess bunchers around the kink point are those that could have earned less if the kink does not exist. It is possible that the introduction of the kink works as a salient reference value for people to adjust their earnings. Although this alternative hypothesis does not seem to be so likely as the null hypothesis, it is an empirical question to examine whether it is true. The idea to examine this key assumption is: if excess bunchers come from those that could have slightly higher income than the kink points, as the bunching theory predicts, then we should see the bunchers at kinks have a lower income growth rate than the nearby non-bunchers. Otherwise, we would see bunchers at kinks have a same or a higher income growth rate than nearby non-bunchers.

In accordance with the bunching estimates above, we focus on pre-reform kinks at 20,000 RMB taxable income and post-reform kinks at 9,000 RMB and 35,000 RMB. Figure 3 shows clear evidence that wage growth rates in the bunching region associated with post-reform kinks at 9,000 RMB and 35,000 RMB are lower than those in neighboring non-bunching area, though the evidence is less clear for the pre-reform 20,000 RMB kink possibly due to less observations. The solid line indicates the kink point and dashed lines embrace the excluded region, where we expect to see a lower income growth rate if the assumption of the bunching approach is correct. Note that since we have excluded round number taxable income observations and only use decimal taxable income observations, the lower wage growth rates

[30]Like all previous bunching papers, we find choosing different bin widths generates only slightly different estimates. To save space, we do not report these results.

in the excluded region are not due to income inertia for those locating at natural focal points. Using the following specification, we statistically test whether wages falling in the excluded region grow significantly lower than wages nearby:

$$Wage\,growth\,rate_{it} = \sum_{j=0}^{q} \beta_j \cdot (z_{it})^j + \gamma \cdot \mathbf{1}_{\boldsymbol{z_{it} \in (z_L,\, z_U)}} + \varepsilon_{it},$$

where z_{it} is the taxable income of individual i at month t relative to kink K, $Wage\,growth\,rate_{it}$ is the wage growth rate (from last month to current month) for individual i from month $t-1$ to month t, $\mathbf{1}_{\boldsymbol{z_{it} \in (z_L,\, z_U)}}$ indicates whether the taxable income z_{it} falls in the excluded region (z_L, z_U). We still take the seventh-degree polynomial, i.e. $q = 7$. The solid smooth lines in figure 3 depict the estimated counterfactual distribution omitting observations in the excluded region. Table 2 shows the test results, where only estimates of γ are reported. For each kink point, the estimate of γ is negative. The estimates are statistically significant at 1% level for post-reform kinks.

Table 2: Do wages in excluded region grow slower than wages nearby?

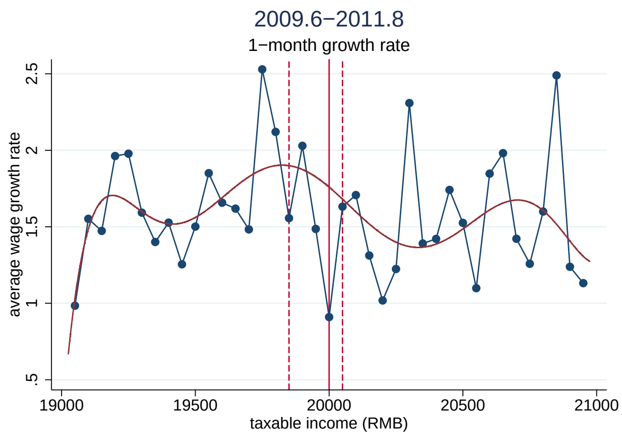
Dep. Var.: wage growth rate	[1]	[2]	[3]
Kink point:	pre-reform taxable income 20000 RMB	post-reform taxable income 9000 RMB	post-reform taxable income 35000 RMB
excluded region	-0.462**	-0.141***	-0.574***
	(0.195)	(0.036)	(0.125)
Observations	3,747	57,696	20,983
R-squared	0.004	0.001	0.024

Notes: Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

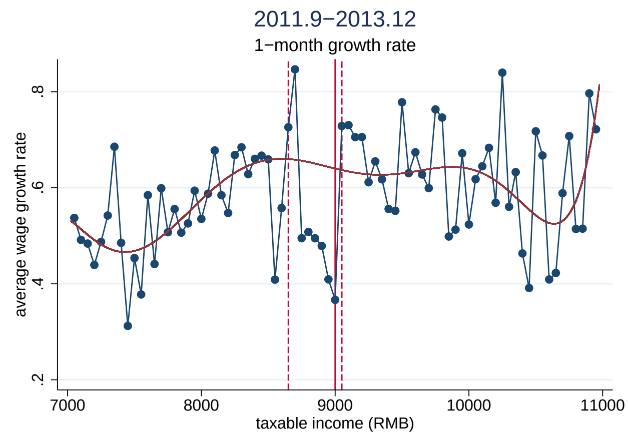
Are bunchers different from nearby non-bunchers? Table A5 compares personal characteristics (i.e. sex, age, occupation, position) for bunchers and nearby non-bunchers for each taxable income kink. Bunchers are defined as the individuals within the excluded region in the above figures, while the nearby non-bunchers are those outside the excluded region but also in the above figures. Since we pool monthly observations, it is possible that a person in some months falls into bunching region while in other months falls into non-bunching region. To address this problem, we define a person as a buncher if he/she ever falls into the excluded region and define a person as non-buncher if he/she never falls into the excluded region.

Figure 3: Wage growth rates around taxable income kinks

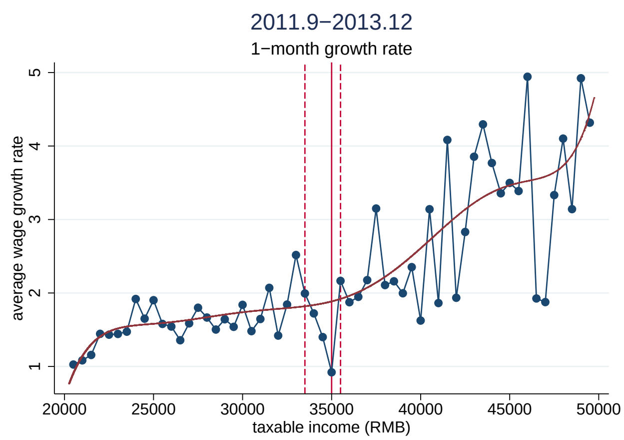
(a) Pre-reform TI kink 20000 RMB



(b) Post-reform TI kink 9000 RMB



(c) Post-reform TI kink 35000 RMB



Notes: The solid smooth curve depicts the estimated counterfactual distribution omitting the obervations in the excluded region, as specified by the area between the dashed lines.

The main finding is that the bunchers at middle-high taxable income kinks tend to have higher position than the non-bunchers, i.e. bunchers are more likely to own some managing power in the workplace (with position of middle deputy and above). This makes sense as those with managing power should have more flexibility to manipulate wage income than the general staff and other people. The occupational comparison does not render sharp contrast and we do not draw clear conclusion from it. Indeed, within each occupation, the heterogeneity in position may be more important in deciding who bunches. In other personal characteristics, bunchers are slightly more likely to be male than female, and there is no systematic difference in age between bunchers and non-bunchers.

4 Tax reform ETI versus bunching ETI

How do the ETI estimates obtained from two standard approaches compare to each other? A preliminary comparison suggests that the tax reform ETI appears to be systematically larger than the bunching ETI, though we show that the two standard approaches do not generate directly comparable estimates. We then develop an approach to allow the two approaches render more directly comparable ETI estimates. After that, we explore how the ETI estimates differ between two approaches and why.

4.1 A preliminary comparison of the ETI estimates obtained from two standard approaches

To start with, we summarize the ETI estimates using two standard approaches in the same country from existing literature. Table 3 shows such a comparison for three countries, China, US, and Denmark. For China, the tax reform ETI is 2.423, much larger than the bunching ETI estimates at the kinks with clear bunching. Using the tax reform approach, Saez et al. (2012) find that for the US, "while there are no truly convincing estimates of the long-run elasticity, the best available estimates range from 0.12 to 0.40"; Kleven and Schultz (2014) find the ETI estimates range from 0.04 to 0.06 for wage earners (and 0.10 for self-employed individuals) in Denmark. Using the bunching approach, Saez (2010) find that for the US, the elasticity is around 0 for wage earners;[31] Chetty et al. (2011) reveal an observed elasticity below 0.02 for wage earnings in Denmark. These empirical findings seem to suggest that for the same country, the tax reform approach generally render higher ETI than the bunching ETI.[32] [33]

But are the estimates from two standard approaches directly comparable? No. The standard tax reform approach differs from the standard bunching approach in two major aspects: time and scope.

[31]The elasticity is around 1 for the self-employed at the first EITC kink, and is around 0.2 for all individuals at the first federal income tax kink (i.e. taxable income \$0).

[32]For other countries that do not have ETI estimates from both approaches, the bunching ETI estimates generally seem quite small. For example, for Sweden, Bastani and Selin (2014) obtain elasticity of 0.001 for all wage earners at the first central government kink during 1999-2005. For Pakistan (during 2006-2009), Kleven and Waseem (2013) use the bunching at notch approach and reveal structural elasticity around 0.03 for wage earners in middle income tax brackets (ranged from 0.03 to 0.28 for self-employed individuals).

[33]Relatedly, Fack and Landais (2016) find that bunching estimates of elasticities of reported charitable contributions with respect to the price of contributions (i.e. one minus the marginal subsidy rate) in France are much smaller than those obtained using the tax reform approach.

Table 3: ETI estimates using two standard approaches

Tax reform approach:	
China (this paper):	2.423 (0.188)
U.S.(Saez et al., 2012):	best available estimates 0.12-0.40
Denmark (Kleven and Schultz, 2014):	wage earners: 0.04-0.06
Bunching approach:	
China: pre-reform kink 20000 RMB	0.10 (0.02)
China: post-reform kink 9000 RMB	0.09 (0.03)
China: post-reform kink 35000 RMB	0.41 (0.07)
U.S.(Saez, 2010):	wage earners: around 0
Denmark (Chetty et al., 2011):	wage earners at top bracket: 0.02

Notes: Standard errors are in parentheses.

Time. The first-differencing specification of the tax reform approach determines the time property of its ETI estimates. For a 1-year differencing specification, the ETI reflects a 1-year response; for a 3-year differencing, the ETI reflects a 3-year response. By contrast, the bunching approach only requires cross-sectional data and it is generally unclear what time property the bunching ETI is capturing. If the bunching ETI is obtained using observations at some arbitrary time point, the bunching ETI only reflects the cross-sectional distribution of taxable income at that time point, thus embedding no time property if no other information is known. By pooling observations in multiple time points together, we obtain an ETI with an even unclearer time property. Thus, the standard tax reform ETI and the bunching ETI normally capture behavioral responses at different time lengths.

Scope. The tax reform approach uses the overall sample and obtains an average ETI for all people. In this sense, the tax reform ETI has a global property. Note that in some specific case, the tax reform approach could also obtain the ETI of some specific group. For example, if a tax reform only reduces the MTR for the top earners while keeping MTRs for other people unchanged, then the tax reform approach would obtain an ETI largely for the top earners. But in general, a tax reform would change the MTR in a broad range of income levels. This is the case for Chinese personal income tax, and in this general case the tax reform ETI has a global property. By contrast, the bunching ETI is estimated based on a close neighborhood of a kink. In this sense, bunching ETI has a local property. A local ETI is hardly comparable to a global ETI.

4.2 Towards a more direct comparison of two approaches

Given the above important differences, the ETI estimates obtained from two standard approaches are not directly comparable. To make the tax reform approach and the bunching approach render more directly comparable estimates, we need to ensure that they reflect the same time property and measure the ETI at the same scope. A two-step framework is devised to do this. First, we develop a revised version for each approach to ensure the ETI estimates from two approaches reflect income responses at the same time length. Second, we ensure the ETI estimates obtained from two approaches reflect the same scope. Since the tax reform ETI has a global property, we try to reveal a global ETI from the bunching ETI estimates in various kinks.[34] We follow the idea in literature (e.g. Chetty (2012), Gelber et al. (2015)) to assume a common underlying ETI for all people, which, with the existence of a fixed adjustment cost, would be consistent with the different observed ETI estimates at different kinks. Using the observed ETI estimates at various kinks, we can reveal the global bunching ETI underlying the observed bunching ETI estimates. This is then comparable to the global ETI obtained from the tax reform approach.

4.3 Revised tax reform approach

The idea of the revised tax reform approach is to estimate an ETI for each period after the tax reform. In particular, we exploit the advantage of the individual monthly panel data to estimate the ETI according to the deviation of the post-reform income growth trend from the pre-reform trend. Instead of using a first-difference specification, the regression specification is like an event-study form:[35] [36]

[34]An alternative way to make this comparison is to use the tax reform approach to estimate the ETI around each kink and compare it with the bunching ETI at that kink. However, using the tax reform approach to estimate ETI for a subsample of individuals whose predicted income is around a kink is dubious in methodology, because for a narrow income interval around a kink, most people would face the same MTR change, which makes identification difficult, if not impossible.

[35]In our specification, the change in the post-reform ln(1-MTR) to the pre-reform ln(1-MTR) is implicitly defined, noting that the individual specific pre-reform ln(1-MTR), e.g. the average ln(1-MTR) of an individual prior to the reform, has been absorbed by the individual fixed-effects.

[36]Admittedly, the average MTR increases over time, even before the tax reform, due to natural income growth and the progressive rate structure. But this tax rate change is much smaller than that induced by the tax reform. Applying the standard tax reform approach (1-year difference specification) using the pre-reform data, with the change in ln(1-MTR) instrumented by the tax rate change driven only by the natural income growth, an approach similar to the "bracket creep" approach by Saez (2003), we find small and insignificant ETI, possibly due to small variation in the MTR caused by the natural income growth and the progressive rate structure. Therefore, we ignore the MTR changes caused by the natural income growth and focus only on the MTR changes caused by the tax reform.

$$log(z_{it}) = \alpha + \sum_{j=1}^{\infty} \beta_j \cdot Post_j \cdot ln(1-\tau_{it}) + \lambda_t + v_i + base\,income\,decile_i \cdot t + \varepsilon_{it}. \quad (3)$$

z_{it} is the taxable income (both pre-reform and post-reform samples are adjusted by the post-reform tax schedule so as to face the same tax base), $Post_j$ is the jth period after the reform, $Post_j \cdot ln(1-\tau_{it})$ is instrumented using $Post_j \cdot ln(1-\tau_i^p)$, τ_i^p is the predicted MTR of individual i with base taxable income (3-month average taxable income prior to the announcement of the tax reform) under the post-reform tax schedule.[37] Monthly fixed effects λ_t control for time-specific shocks to z_{it} that are common across individuals. Individual fixed-effects v_i control for variation in outcomes across individuals that are constant over time. To account for the heterogeneous income growth trends for different income levels, we include a set of flexible controls $base\,income\,decile_i \cdot t$, i.e. the interaction between base income decile dummies and a linear time trend.[38] β_j is the ETI in the jth period after the reform. We use a 3-month period to mitigate monthly fluctuations and also for a direct comparison to the 3-month dynamic bunching ETI estimates below. Since the tax reform proposal was passed on June 30 and was put into effect on September 1st, 2011, people might respond as early as two months prior to the implementation of the reform. Thus, we exclude July and August of 2011 from the regression sample.

Our specification assumes that, prior to the tax reform, the average $log(z_{it})$ increases linearly over time, as is confirmed by figure 4. The tax reform changed the MTR for some people permanently, which drives the dynamic adjustments of their taxable income. Using the pre-reform trend as the benchmark, we can estimate the ETI for each post-reform period by tracking the deviation of $log(z_{it})$ from its pre-reform trend. Our specification also greatly avoids the mean reversion issue, which could be acute when using the first-difference specification, since we use a long pre-reform trend of income rather than income at a time point as the benchmark.

Before showing the ETI estimates, we provide graphic evidence on the evolution of the MTR and the log of taxable income around the tax reform. The graphical evidence is very helpful to judge whether the dynamic ETI estimates obtained below are reasonable. Figure 4 shows clearly that people facing an MTR increase experienced lower income growth after the reform, while those facing an MTR decrease had much faster income growth after the

[37] We have tried to produce an estimate with just $Post \cdot ln(1-\tau_{it})$ and obtain an ETI estimate of 1.8. But this estimate has no clear time property, as it is an average ETI of all post periods. Thus, it is not directly comparable to our standard tax reform ETI estimate, which can be interpreted as a one-year ETI due to its first-difference specification using one-year lags.

[38] Using the interaction between base income percentile dummies and a linear time trend renders very similar results.

reform.[39] The figure implies that for the MTR increase group, the dynamic ETI increases over time, while for the MTR decrease group, the ETI experiences dramatic increase in initial periods after the reform, and stay largely constant in the end. This reflect asymmetric income responses to MTR increases versus decreases. Under our setting, MTR increases happen at high income levels while MTR decreases happen at low income levels. Therefore, these asymmetric income responses may be due to different capabilities to adjust income for high earners versus earners.[40]

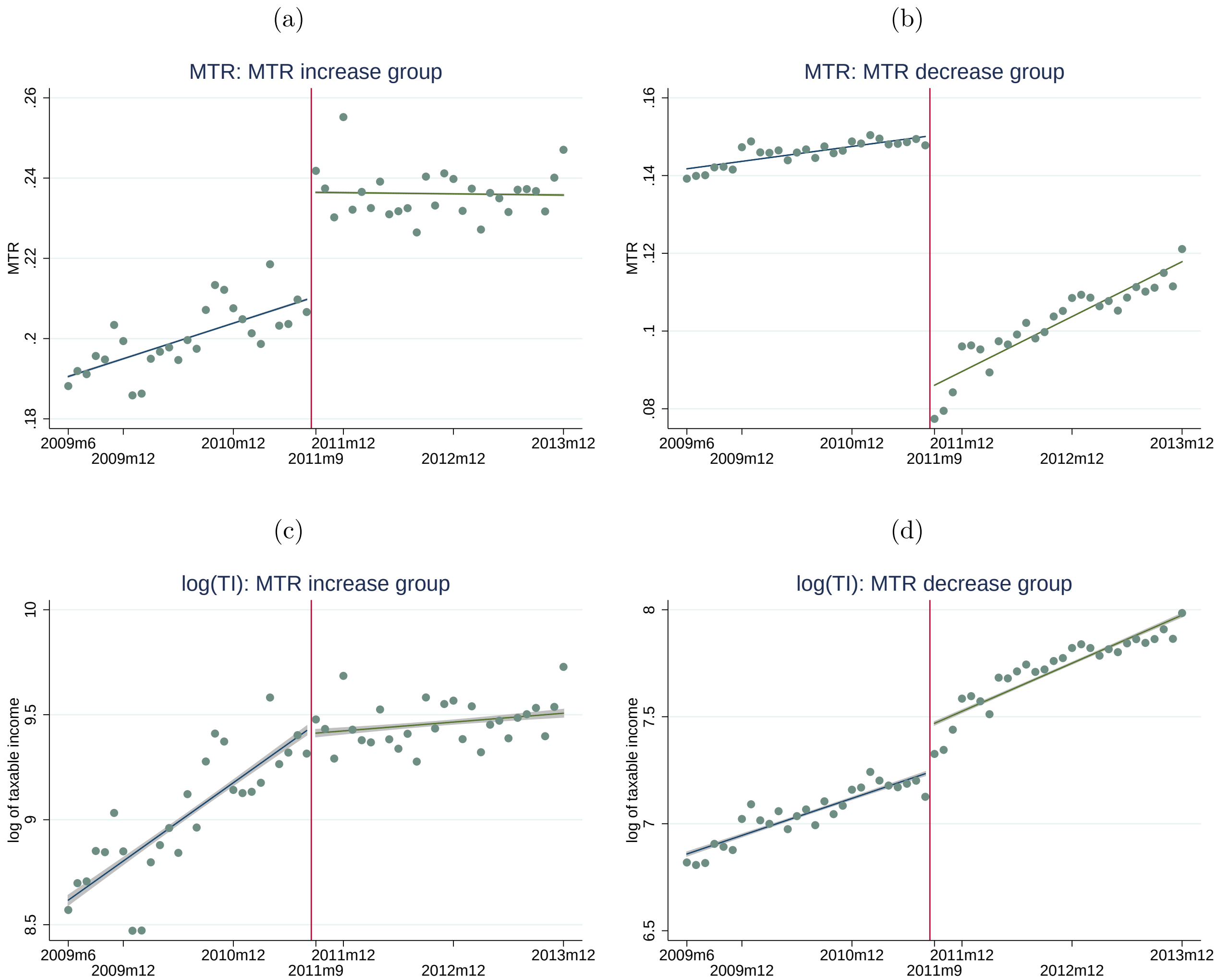
With an anticipation of a future MTR change, besides the labor supply adjustment, people might also respond in other margins. One potential response is shifting income intertemporally. For example, anticipating an MTR decrease in a forthcoming tax reform, people would have incentive to shift income from pre-reform to post-reform periods. If such an intertemporal income shifting is salient, we should see people facing MTR decreases exhibit a dip in income between the annoucement of the tax reform and its implementation (July and August, 2011), and a surge in income in months after the reform. For those anticipating an MTR increase, we should see the opposite. This is observed in Denmark for wage income by Kreiner et al. (2016). However, in figure 4, we see no evidence for intertemporal income shifting. This cannot be explained by a smaller change in the MTR as compared to that in Denmark.[41] One possible explanation is the much shorter periods to make adjustments between the announcement and the implementation of the reform in China. China's 2011 PIT reform proposal was passed on June 30 and put into effect on September 1st, 2011, leaving only two months prior to the tax reform to make adjustment. By contrast, the Danish reform was passed in parliament at the end of May 2009 and changed the tax scheme from 2010 onward, leaving seven months to take adjustments. Kreiner et al. (2016) show that the income adjustment was obvious only starting from November, suggesting sufficient time is needed to make such adjustments.

[39]We judge whether an individual faces an MTR increase or decrease based on the baseline taxable income, which is the average taxable income throughout April-June, 2011. From figure A2, it is clear which people face an MTR increase, which face an MTR decrease. For people who do not face an immediate MTR change at the reform month, most of them would face an MTR increase later due to the natural wage growth and the progressive tax rate schedule.

[40]It would be interesting to examine whether similar people (i.e. with similar incomes) respond asymmetrically to tax increases versus decreases. Relatedly, Benzarti et al. (2017) find that prices respond more to increases than to decreases in value-added tax rates.

[41] Kreiner et al. (2016) study the tax reform that reduced the highest marginal tax rate on earnings from 63 percent to 56 percent. Figure 4 shows that the average MTR decreases from 15 percent to 9 percent for those people facing MTR decreases, and the average MTR increases from 21 percent to 24 percent for those facing MTR increases, which is comparable to the tax rate change in Denmark.

Figure 4: Evolution of the marginal tax rate and the taxable income

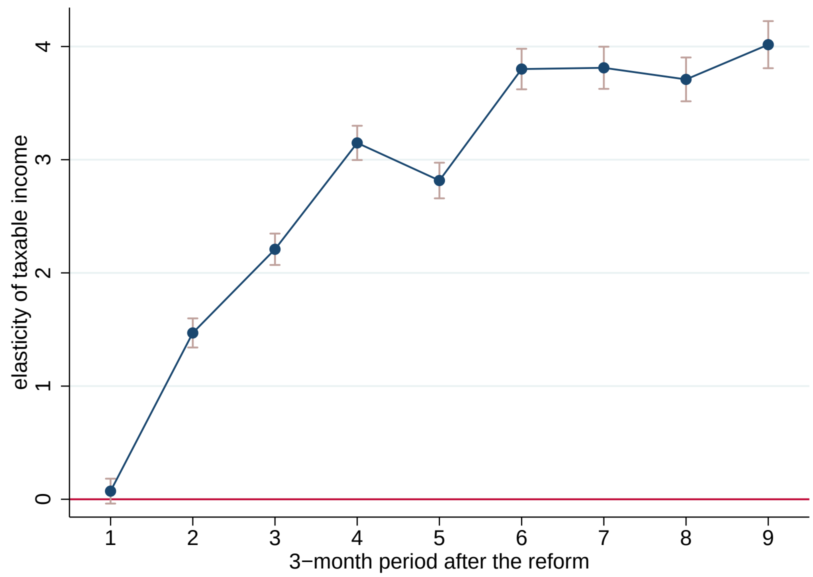


Notes: A linear polynomial is fitted with 95% CI shown in the figure. Each dot represents the average value in a given month.

Figure 5 shows clearly that the tax reform ETI increases steadily over time within the first year after the reform, and then converges to around 4, consistent with the income trends in figure 4.[42] Since a much larger portion of people (lower wage earners, over 70%) experience MTR decreases while a relatively small portion of people (higher wage earners, less than 10%) experience MTR increases, the ETI estimates reflect more of the income trend of the MTR decrease group. The concavely increasing trend of the ETI estimates shows that income adjusts gradually after a tax reform. The larger longer-run ETI also suggests a larger

[42]The magnitue of the ETI estimate is also in line with the naive ETI implied by figure 4. For example, the average log taxable income of the MTR decrease group increases by around 0.25 in the long run, and the average MTR decreases from about 0.15 to about 0.09 after the tax reform. This implies a long-run ETI of 3.66 (=0.25/(ln(1-0.09)-ln(1-0.15))), which is close to our estimate of 4.

Figure 5: ETI estimates over time: revised tax reform approach



Notes: A 95% CI of the estimates is shown in the figure.

efficiency loss of tax in the long run. For our sample city of China, this figure suggests that the long-run ETI is about 4. In addition, the ETI estimate after one year of the tax reform is around 3, close to our ETI estimates using the standard tax reform approach with one-year difference specification.

Potential bias caused by kinks. There is some concern that the existence of kinks might bias the tax reform ETI estimates. Due to the natural income growth, when one's income grows closely to the next kink, his income growth rate might slow down according to the bunching theory. For people experiencing MTR increases over a kink, the ETI may be under-estimated, since people won't reduce their income below the kink; for people experiencing MTR decreases, the ETI may be over-estimated, since people originally hindered by the kink will add to the aggregate responses of those not around the kink. Overall, the direction of the aggregate bias is unclear. However, since income growth is unlikely to be hampered by the kink for a long time (shown below), the bias should be very small. In fact, excluding observations falling in the bunching regions in figure 2 does not change our ETI estimates, which is shown clearly in figure A9. Moreover, excluding observations around all kinks only slightly changes the estimates, as also found in Kleven and Schultz (2014).[43] Table A6 further shows how kinks could bias the tax reform ETI estimates. For people whose pre-reform TI is around the pre-reform kinks (i.e. falling in the bunching region of kinks),

[43]For non-bunching kinks, we exclude the ± 100 RMB inverval around each kink.

their ETI estimates are significantly larger; while for people whose predicted post-reform TI (based on pre-reform TI and income growth rates) is around the post-reform kinks, their ETI estimates are significantly smaller. These are consistent with the predictions by the bunching theory, as anticipating to be around kinks lowers the income growth rates, while getting rid of the constraint of kinks has income growth faster than normal. Finally and importantly, the ETI estimates for people not affected by kinks are very close to our original estimates, again showing kinks have little impact on the tax reform ETI estimates using the full sample.

4.4 Revised bunching approach

To come up with the global bunching ETI comparable with the dynamic tax reform ETI estimates, we use a two-step approach. First, we explore the dynamics of bunching behavior in response to the kink changes due to the tax reform. This will render a dynamic sequence of ETI estimates for each kink. Second, for each time period, using the approach described in detail below, we estimate a common structural ETI underlying the observed bunching ETI estimates at various kinks. This two-step approach finally gives us a sequence of global structural ETI for each period after the reform. In the following, we first introduce the approach to reveal the global underlying ETI using observed ETI estimates at various kinks. Then we apply this approach to reveal the global underlying ETI based on the observed ETI estimates using the standard bunching approach. In this process, we show that the global underlying ETI is not only consistent with the observed bunching at middle-high kinks, but can also explain the non-bunching evidence at lower kinks. Then we apply this approach to obtain a sequence of global bunching ETI estimates and compare them with those obtained from the revised tax reform approach.

4.4.1 Bunching approach with optimization frictions: conceptual framework

The standard bunching approach renders an observed ETI for each kink by assuming no optimization frictions. However, such an assumption seems unrealistic in general. Whenever possible, it is ideal to explicitly estimate the size of the optimization frictions to see if it can be reasonably ignored. We follow the literature to call the standard bunching ETI estimate $\widehat{\varepsilon}$ the observed elasticity, and we want to know what underlying elasticity ε may have driven these observed elasticities when there exist optimization frictions. Here optimization frictions are defined broadly as any forces (e.g. real costs of adjusting earnings, inattention, or information costs) that prevent people from re-achieving optima. We adopt a modified approach based on Gelber et al. (2015) and use adjustment costs interchangeably with optimization frictions. We assume that there is a fixed ETI ε and a fixed utility cost ϕ that must be paid whenever adjusting income, which changes the utility function to $u(z-T(z), \frac{z}{n}) - \phi$ if an adjustment is made.[44] To estimate ε and ϕ, at least two empirical moments are required. A key difference between our approaches is that Gelber et al. (2015) rely on a before-and-after comparison of bunching at the same kink where the jump in marginal tax rates reduced, while we use bunching at different kink points to construct empirical moments. ε and ϕ are assumed fixed across kinks, which can be interpreted as the average elasticity and average adjustment cost for all people.

Consider a two-tier budget line with a kink at taxable income K. The MTR is τ_1 below K and increases to $\tau_2 = \tau_1 + \triangle\tau$ above K. When there is no adjustment cost, all individuals initially choosing $(K,\, K + \triangle z)$ would adjust income to K, and the person initially choosing $K + \triangle z$ is called the marginal buncher because he is indifferent between adjusting and not. When there is a positive adjustment cost ϕ, however, individuals initially choosing $(K,\, K + \underline{z})$ would find it too costly to adjust income to K, where $\underline{z}$ depends on ϕ. We assume $\underline{z} < \triangle z$ because otherwise the adjustment cost would be too large that there would be no bunching at the kink at all. Now all individuals initially choosing $(K + \underline{z},\, K + \triangle z)$ would adjust income to K.[45]

Then the excess bunching with adjustment cost is

$$B = \int_{K+\underline{z}}^{K+\triangle z} h_0(z)dz \simeq h_0(K) \cdot (\triangle z - \underline{z}) = \frac{h_0^W(K) \cdot (\triangle z - \underline{z})}{W},$$

where the second equality holds approximately when $\triangle z$ is small or when the density is

[44] Although the optimization frictions could be very large, as estimated in Kleven and Waseem (2013) in the notch case, the original approach by Saez (2010) does not consider the optimization frictions and thus cannot uncover the underlying structural elasticity. To obtain the underlying elasticity, two approaches are considered in literature (see Kleven (2016) for a detailed review). One approach is to utilize the variation in the size of kinks that is orthogonal to the underlying elasticity and optimization fractions. Along this approach, Gelber et al. (2015) assume a fixed optimization friction and underlying elasticity across kinks, which ensures them to be identified with two kinks. For more discussion on the assumption of a fixed adjustment cost, see Gelber et al. (2015). By contrast, Chetty et al. (2010) consider a more involved model where there are more parameters and thus need more bunching moments to make them fully identified. Another alternative approach to estimate the structural ETI from observed ETI estimates is the bound approach by Chetty (2012). The advantage of this bound approach is that it does not make specific assumptions on the utility function. But the cost of imposing weaker assumptions is that it requires large tax reforms to obtain tight bounds, which is not well satisfied in our case. In Online Appendix E we apply this approach and show very wide bounds. Thus, in our case, the bound estimates are not very informative and we stick to the parametric approach to uncover the structural elasticity.

[45] We follow Gelber et al. (2015) to make an intuitive assumption that the benefit of adjusting income to the kink K is increasing in distance from the kink for initial earnings in the range of $(K,\, K + \triangle z)$. This assumption is true, for example, when utility function is quasi-linear. See their paper for a more detailed argument.

uniform in interval $(K,\ K+\triangle z)$. Then we have

$$b \equiv \frac{B}{h_0^W(K)} \simeq \frac{\triangle z - \underline{z}}{W}. \tag{4}$$

To estimate ε and ϕ, we take a parametric approach and assume a quasi-linear utility function $u(c,z;n) = c - \frac{n}{1+\frac{1}{\varepsilon}}(\frac{z}{n})^{1+\frac{1}{\varepsilon}}$, following Saez (2010), Kleven and Waseem (2013), and Gelber et al. (2015). Due to this assumption, ϕ has the same money metric unit as consumption. With a linear tax schedule $T(z) = \tau z$, we have $c = z - \tau z$ and the utility maximization renders the optimal choice of earnings $z = n(1-\tau)^{\varepsilon}$. Then the person initially choosing $K + \underline{z}$ has ability $\underline{n} = \frac{K+\underline{z}}{(1-\tau_1)^{\varepsilon}}$. Since this person is indifferent between adjusting income to K and making no adjustment, we have

$$u(K,\ \tau_1;\ \underline{n}) - \phi = u(K+\underline{z},\ \tau_2;\ \underline{n}),$$

which renders

$$K(1-\tau_1) - \frac{\underline{n}}{1+\frac{1}{\varepsilon}}(\frac{K}{\underline{n}})^{1+\frac{1}{\varepsilon}} - \phi = K(1-\tau_1) + \underline{z}(1-\tau_2) - \frac{\underline{n}}{1+\frac{1}{\varepsilon}}(\frac{K+\underline{z}}{\underline{n}})^{1+\frac{1}{\varepsilon}}.$$

Plugging $\underline{n} = \frac{K+\underline{z}}{(1-\tau_1)^{\varepsilon}}$ to the above equation and simplify it, we obtain

$$\frac{(K+\underline{z})(1-\tau_1)}{1+\frac{1}{\varepsilon}}[1-(\frac{K}{K+\underline{z}})^{1+\frac{1}{\varepsilon}}] - \phi = \underline{z}(1-\tau_2). \tag{5}$$

By definition, the underlying elasticity is $\varepsilon = \frac{\triangle z/K}{\triangle\tau/(1-\tau_1)}$, which combined with (4) renders

$$b \simeq \frac{\varepsilon \cdot \frac{\triangle\tau}{1-\tau_1} \cdot K - \underline{z}}{W}. \tag{6}$$

From (5) and (6), we can obtain an implicit relation $b = b(\varepsilon, \phi)$ at kink K. Denote by b_i the theoretical bunching at kink K_i and by $\widehat{b_i}$ the empirical moments estimated using (2).

We employ minimum distance estimation to estimate (ε, ϕ). The idea is to seek the values of the parameters that make theoretical bunchings b_is as close to the empirical bunching $\widehat{b_i}$s as possible:

$$(\widehat{\varepsilon}, \widehat{\phi}) = argmin_{(\varepsilon,\phi)} \sum_i (b_i - \widehat{b_i})^2.$$

Following Gelber et al. (2015), we obtain our estimates by minimizing the above equation numerically. Solving this problem requires evaluating b_i at each trial guess value of (ε, ϕ).

Here, we use observed bunching at various kinks to form empirical moments to estimate the two parameters $(\widehat{\varepsilon}, \widehat{\phi})$. Since $(\widehat{\varepsilon}, \widehat{\phi})$ is a function of the estimated amount of bunching $\widehat{b_i}$s, we can estimate their standard errors using the bootstrapped estimates of $\widehat{b_i}$s. In particular, we can solve for $(\widehat{\varepsilon}, \widehat{\phi})_j$ for the jth draw of $\widehat{b_i}$s, and the standard deviations of $(\widehat{\varepsilon}, \widehat{\phi})_j$ are the bootstrapped standard errors of $(\widehat{\varepsilon}, \widehat{\phi})$. With the estimates of $(\widehat{\varepsilon}, \widehat{\phi})$, we can also plug them back to (5) and (4) to obtain the estimates for $\underline{z}$ and $\triangle z$ for each kink K. The standard errors of $\underline{z}$ and $\triangle z$ are obtained similarly to those of $(\widehat{\varepsilon}, \widehat{\phi})$.

4.4.2 Revealing a common structural elasticity using standard bunching ETI estimates

To estimate both the underlying elasticity ε and the fixed adjustment cost ϕ, we need at least two empirical moments. In our standard bunching estimates (without optimization frictions), we focus on pre-reform taxable income kink at 20,000 RMB and post-reform kinks at 9,000 RMB and 35,000 RMB due to clear bunching evidence at these places. We thus use the bunching estimates at these kinks to form three empirical moments. Table 4 shows the estimates. If we assume a constant underlying elasticity of taxable income ε and a constant adjustment cost ϕ at all the three kinks, our estimates suggest that ε is 0.14 and ϕ is 1.22 RMB. Although the adjustment cost is small, it precludes those initially earning within around 30 RMB above the kink from adjusting earnings, while the marginal bunchers vary from earning 76 RMB above the kink to 317 RMB above the kink for the three kinks.

A key point here is that a small optimization friction could preclude a non-negligible amount of people from adjusting earnings to the optima. This could help explain why we observe close to zero bunching at lower taxable income kinks. To see this, we explore the implication of our estimates of ε and ϕ based on bunching kinks for the non-bunching kinks. With the estimate of ε, we can use the definition of ε to obtain $\triangle z = \frac{\varepsilon \cdot K \cdot \triangle \tau}{1-\tau_1}$ for each non-bunching kink. Then for each non-bunching kink, we calculate the optimization friction $\underline{\phi}$ that would make the marginal buncher without optimization friction (i.e. the person originally choosing $K + \triangle z$) indifferent between adjusting income to K or not.[46] That is, $\underline{\phi} = \frac{(K+\triangle z)(1-\tau_1)}{1+\frac{1}{\varepsilon}}[1-(\frac{K}{K+\triangle z})^{1+\frac{1}{\varepsilon}}] - \triangle z(1-\tau_2)$. Table 4 shows that for each non-bunching kink except for the largest non-bunching kink (post-reform kink 4,500 RMB), $\underline{\phi}$ is smaller than ϕ, suggesting that the small estimated adjustment cost preclude any bunching at the lower kinks. As for the post-reform kink 4,500 RMB, figure A4 actually show small scale bunching, though not salient enough for our estimation purpose. Thus, a common structural elasticity

[46] We do not insert the estimates of ε and ϕ into equation 5 to solve for $\underline{z}$ for each kink because it does not necessarily have a solution.

Table 4: Estimates for underlying ETI ε when there is fixed adjustment cost ϕ

	parameters	taxable income kink point K	point estimate	bootstrapped standard error
	underlying ETI ε	assumed constant across all kinks	0.14	0.02
	adjustment cost ϕ (RMB)		1.22	0.61
Bunching taxable income kinks	Δz (RMB)	pre-reform kink 20000 RMB	169.75	28.64
		post-reform kink 9000 RMB	76.39	12.89
		post-reform kink 35000 RMB	316.87	53.46
	z (RMB)	pre-reform kink 20000 RMB	26.83	13.90
		post-reform kink 9000 RMB	32.64	18.68
		post-reform kink 35000 RMB	25.58	13.01
Non-bunching taxable income kinks	Δz (RMB)	pre-reform kink 500 RMB	3.57	0.60
		pre-reform kink 2000 RMB	15.09	2.55
		pre-reform kink 5000 RMB	39.94	6.74
		post-reform kink 1500 RMB	14.70	2.48
		post-reform kink 4500 RMB	67.90	11.46
	adjustment cost ϕ (RMB)	pre-reform kink 500 RMB	0.09	0.02
		pre-reform kink 2000 RMB	0.39	0.07
		pre-reform kink 5000 RMB	1.02	0.17
		post-reform kink 1500 RMB	0.53	0.09
		post-reform kink 4500 RMB	3.55	0.60

Notes: When there is fixed adjustment cost ϕ, for a kink point K, those initially choosing between $K + \underline{z}$ and $K + \triangle z$ will adjust income to K. Adjustment cost $\underline{\phi}$ makes the marginal buncher indifferent between adjusting income to K or not.

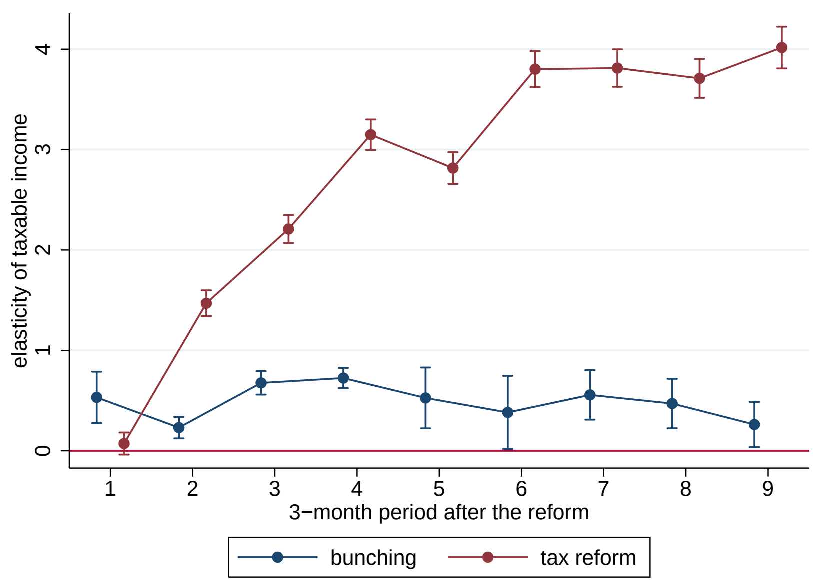
with small adjustment cost could well explain both the observed bunching at middle-high TI kinks and the lack of bunching at lower kinks.

4.4.3 Estimating the global bunching ETI over time

The success of the structural ETI with optimization frictions to explain the bunching at middle-high taxable income kinks and the non-bunching at lower kinks provides support for assuming a common structural ETI underlying the observed bunching at various kinks. But the above bunching estimates are obtained by pooling all pre-reform (or post-reform) observations together and do not have a clear time property. To obtain the bunching ETI estimates over time, we explore dynamic bunching responses to changes of kinks because of the 2011 tax reform. In particular, we focus on the introduction of new kinks (9,000 RMB & 35,000 RMB) and explore the dynamic evolution of bunching scale after the reform. We do not use the abolition of old kinks (e.g. 20,000 RMB) due to asymmetric bunching responses to new kinks and to old kinks. It is observed in our data that bunching immediately disappear after the old kink (20,000 RMB) was abolished.[47] Thus, it seems desirable to just focus on bunching responses to the introduction of new kinks. Fortunately, we need at least two empirical moments to estimate ε and ϕ, and the introduction of kinks at 9,000 RMB & 35,000 RMB provide two such empirical moments. In addition, since 20,000 RMB is included in the range between 9,000 RMB and 35,000 RMB, using these two kinks would not narrow our coverage of income levels.

Since there is no bunching at the new kinks before the reform, we can simply apply the standard bunching approach to estimate the ETI at each kink for each period after the reform. Then for each period, we apply the above approach to uncover the structural ETI. One advantage of such approach is that it could allow optimization frictions to vary flexibly over time. In principle, we can estimate a bunching ETI for each month. But observations at the middle-high kinks are too few for each month, so we combine three months as a period. To make a direct comparison, we show the revised tax reform ETI by 3-month period and overlap it with the dynamic sequence of global bunching ETI in figure 6.[48]

Figure 6: Global ETI estimates over time: tax reform approach vs. bunching approach



Notes: A 95% CI of the estimates is shown in the figure. The bunching ETI are estimated using the decimal sample.

[47]The asymmetric bunching responses to new kinks and to old kinks are also observed in Best and Kleven (2016), under the setting of housing transaction taxes.

[48]Although the global bunching ETI estimates shown in figure 6 are based on the decimal sample, in Online Appendix F we show that using the full sample would generate similar dynamic patterns of bunching ETI estimates.

4.5 Tax reform ETI vs. bunching ETI: key differences and potential explanations

Figure 6 shows sharp contrast between the tax reform ETI and the bunching ETI over time. It clearly documents two key differences:

1. The tax reform ETI estimates increase concavely over time, while the bunching ETI estimates are stable over time.

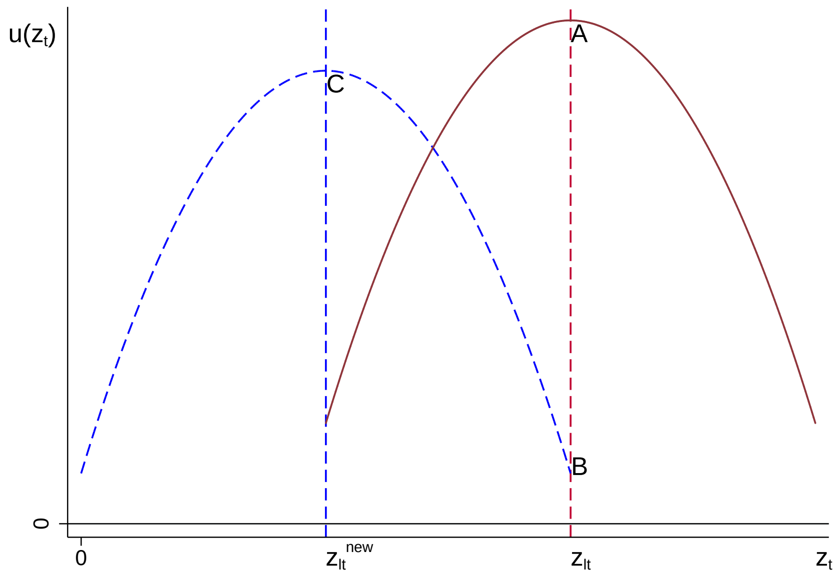
2. The long-run tax reform ETI estimates (around 4) are much larger than the bunching ETI (around 0.5).

If the tax reform approach and the bunching approach both capture the same structural elasticity that governs people's behavior, we should not observe such differences. Thus, the two approaches must have captured very different behavioral responses. What behavioral response does each approach fundamentally capture? Why do they capture such different behavioral responses? In the following, we explore the potential answer to these questions. We try to use a parsimonious model to account for these differences altogether. The key ingredients of the model include: (1) infrequent but permanent adjustments of hours of work with no additional cost, and (2) temporary adjustment of hours of work with some adjustment costs. The model is able to not only reconcile these two key differences, but also generates other testable implications supported by our data.

4.5.1 A simple model

Denote hours of work at time t by h_t and the wage rate by w_t. The latent income (corresponding to the optimal hours of work h_t under current tax rate) is $z_{lt} = h_t w_t$. We assume that w_t is exogenous to employees and grows at a rate g, i.e. $w_t = w_0(1+g)^t$, while h_t can be adjusted permanently at no additional cost, with an exogenous probability $q < 1$ in each period. Here $q < 1$ reflects that due to various restrictions (e.g. constraints of pre-existing contract, time to adjust production organization, and time needed to switch to a new job), people need time to adjust their hours of work to the new optima. Denote taxable income by z_t, which would be the same as z_{lt} without interruption. In addition, we assume that in each period after the potential permanent adjustment of h_t, people can still adjust h_t temporarily by paying an additional cost; this would make z_t deviate from z_{lt} only for the current period. Suppose the adjustment cost takes the form of $\phi + c \cdot f(|z_{lt} - z_t|)$, where with $\phi > 0$, $c > 0$, $f' > 0$, $f'' > 0$. Note that $z_{lt} = h_t w_t$, corresponding to the income under the old optimal hours of work h_t, while $z_t = h_t^{temp} w_t$, corresponding to the income associated with the temporarily adjusted hours of work h_t^{temp}. This assumption imbeds the idea that the costs of making a temporary adjustment of hours not only involve a fixed cost ϕ common to every one (corresponding to the technical cost of adjusting h_t temporarily, as well as search costs and informational costs), but also include a penalty, which convexly increases with the deviation from the hours of work required by the current contract.[49]

Figure 7: The impact of a tax change on the optimal income choice



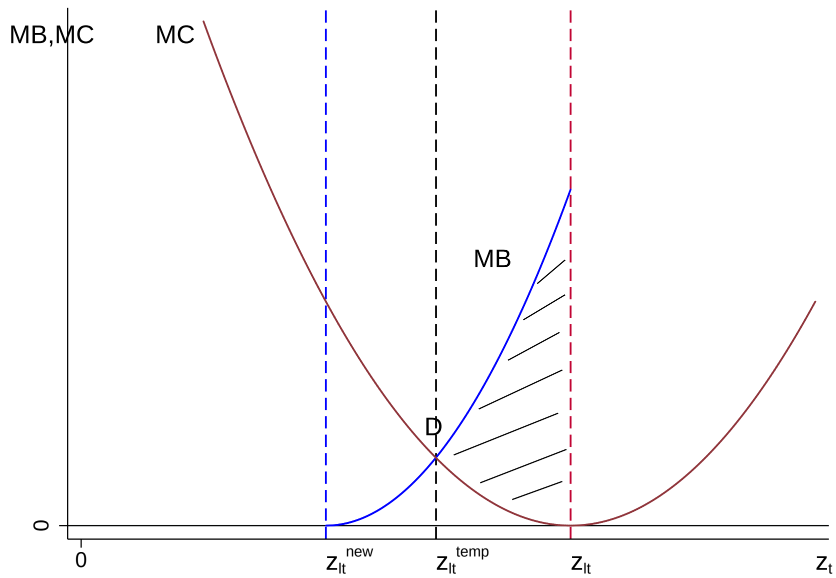
Notes: Without considering adjustment costs, when the MTR increases from τ_1 to τ_2, the utility curve shifts leftwards, and the optimal choice of income z_t (determined by hours of work h_t, for a given wage rate w_t) decreases from z_{lt} to z_{lt}^{new}.

Figure 7 illustrates how a tax change affects the optimal income choice when there is no adjustment cost. When the MTR increases from τ_1 to τ_2, the utility curve $u(z_t)$ (a value function of income z_t) shifts leftwards, and the optimal choice of income z_t (determined by hours of work h_t, for a given wage rate w_t) will decrease from z_{lt} to z_{lt}^{new}. The initial optimal choice lies at point A. When the tax rate increases, if a person does not adjust behavior, his utility would drop to point B; the benefit of adjusting income increases as he moves from B throughout C. But the marginal benefit (MB) of adjusting decreases from B throughout C,

[49]A fixed cost ϕ for making a temporary adjustment of h_t, instead, is not sufficient to explain the stylized facts. Under our model, if there is bunching at a kink, it implies that people slightly above the kink find it worthwhile to making a temporary adjustment. If a fixed cost ϕ applies for all people making a temporary adjustment, then all people with income well above the kink would find it worthwhile making an adjustment. Therefore, they would immediately adjust to the new long-run optima, even when they do not face a chance to permanently adjust h_t. This, however, would be inconsistent with the concavely increasing tax reform ETI estimates.

and would be negative below C. The positive segment of the MB curve is shown in figure 8, as this is the only possible part for a potential adjustment.

Figure 8: Marginal benefit and marginal cost of adjusting income



Notes: The upward sloping line depicts the positive segment of the marginal benefit curve, and the U-shaped curve depicts marginal costs in relevant region.

For the infrequent but permanent adjustment of hours of work, there is no additional adjustment cost. If the tax change is simply an increase of the MTR in all income levels, then all people would simply adjust hours of work to the new optimum when they face such a chance, which corresponds to the adjustment from z_{lt} to z_{lt}^{new}, as illustrated in figure 7. But we are considering an increase of the MTR only for incomes above a kink K. Then there are two cases. For people with z_{lt} well above K, their new optimum z_{lt}^{new} should be higher than K, and therefore they would adjust hours to the new long-run optimum when they face such a chance. For people with z_{lt} close enough to K, who face a smaller z_{lt}^{new} than K, the kink induces them to adjust h_t less sufficiently to the new long-run optimum. The wage growth rate g and the discount rate δ together determine the extent of their adjustment of h_t. A larger g will make the gains from an insufficient adjustment of h_t smaller, since people will then stay close to the kink for a shorter time; a smaller δ (corresponding to a more patient person) will also make the gains from an insufficient adjustment of h_t smaller, since the gains from adjusting income to the kink has a lower weight in the discounted utility of all periods. Therefore, people with a larger g or a smaller δ will adjust more fully to the new long-run optimal hours of work. Finally, there are some people who are not currently experiencing an increase in the MTR yet forecast such a change in the near future as their wage rate grows steadily. For these people, when they face a chance to permanently adjust hours of work, they would first adjust to a weighted average of their short-run optimum (under current tax rate) and their new long-run optimum, and adjust fully to their new long-run optimum when they face next chance to make a permanent adjustment.

In each period, people can temporarily adjust hours of work by paying costs $\phi + c \cdot f(|z_{lt} - z_t|)$. Since we have assumed $c > 0$, $f' > 0$, $f'' > 0$, the marginal cost (MC) of making a temporary adjustment increases convexly with the adjustment distance, which is illustrated in figure 8. Let D be the cross point of MC and MB curves. Starting from z_{lt}, MB decreases and MC increases as people adjust towards z_{lt}^{new}. The gain from the temporary adjustment is the area below MB and above MC. Obviously, the largest gain is obtained when people adjust to z_{lt}^{temp}, the income level corresponding to D. This gain, indicated by the shaded area, should then be compared with the fixed cost ϕ. If the shaded area is smaller than ϕ, then the optimal choice is to make no temporary adjustment. If the shaded area is larger than ϕ, then there are two cases. For people with z_{lt} far above K, their cross point D is likely to be higher than K, their optimal choice is to adjust temporarily to z_{lt}^{temp}. For people with z_{lt} close enough to K, their cross point D is likely to be lower than K; their choice is binded by the kink, and therefore they will temporarily adjust to K.

In the following, we show that the major implications of the above simple model can reconcile the key differences between the ETI estimates obtained using two approaches. Moreover, the model generates other testable implications supported by our data.

4.5.2 Major implications from the model

Concavely increasing tax reform ETI. When people face a chance to permanently adjust h_t, most of them (those well above the kink) will simply adjust to the new long-run optimal hours, some of them (those close enough to the kink) will adjust partially to the new long-run optimum. As long as the wage growth rate g is not too small, or people are patient enough (δ small), people will mostly adjust h_t to the new long-run optima when they face a chance to make a permanent adjustment. Therefore, the permanent adjustment of h_t has little contribution to the bunching at the kink. Since in each period people independently face such a chance w.p. $q < 1$, we should observe they aggregately gradually adjust hours of work over time. This gradual adjustment to the new long-run optima is captured by the tax reform ETI estimates.[50] The fraction of people that has shifted to the new optimum after

[50]The tax reform ETI etsimates also reflect the temporary adjustment of h_t. We have shown that with adjustment costs, people temporarily adjust h_t only partially. When the costs for a temporary adjustment

j periods would largely equal to $1 - (1-q)^j$, yielding a curve roughly corresponding to the concavely increasing ETI estimates obtained from the revised tax reform approach.[51]

Small and stable bunching ETI. As the permanent adjustment of h_t contributes little to the bunching at the kink, bunching mostly reflects the temporary adjustment of h_t, and is therefore ruled by the adjustment cost. Without such cost, all people would immediately adjust to the new long-run optimal hours, and the bunching ETI and the tax reform ETI would both immediately reflect the long-run underlying ETI. When there are convex costs for a temporary adjustment of h_t, however, people will adjust hours only partially relative to the new long-run optima. When the costs for a temporary adjustment are large enough, the bunching ETI can be arbitrarily small. Since in each period, people with income falling into the same certain range will find it worth adjusting income (via adjusting h_t) to the kink (though they are different groups of people in different periods under a normal wage growth rate $g > 0$), assuming a largely stable income distribution, the bunching ETI estimates will be stable over time.

Convergence of two estimators in a stagnant economy. If instead $g = 0$, i.e. an economy is stagnant, the bunching approach is then able to capture even the permanent adjustment of behavior. In this case, we should expect to see a gradual (and concave) increase of bunching at the kink over time. People close enough to the kink but not making a temporary adjustment of h_t (because adjustment costs are higher than the accrued benefits from a temporary adjustment) would gradually face a chance in each period w.p. $q < 1$ to permanently adjust h_t without paying additional cost. Now they can adjust hours so that income is adjusted to the kink. The gradual permanent adjustment implies a concavely increasing pattern of bunching at the kink K. Therefore, our model implies that in a stagnant economy, the bunching ETI will not only capture the temporary adjustment of h_t in the short run, but also reflect the permanent adjustment of h_t in the long run. Thus, the bunching and the tax reform ETI estimators would finally converge to the same underlying parameter. Although this pattern is not ready to test using our data, it can be tested in an otherwise

are large enough, the temporary adjustment of h_t can be small enough, dwarfed by the full adjustment of those facing a chance to parmanently adjust h_t. Therefore, the tax reform ETI estimates mostly reflect the permanent adjustment of h_t.

[51]Our model implies concavely increasing tax reform ETI estimates for people at all income levels, since all people will gradually face the opportunity to permanently adjust hours. This implication is consistent with the empirical evidence in figure 4, which shows that people with lower income levels (those experiencing MTR decreases) experience gradual income adjustments, just like people with higher income levels (those experiencing MTR decreases) do. The lack of bunching at lower kinks and the existence of bunching at higher kinks, however, are explained by the fixed costs (now plus an additional convex cost) for a temporary adjustment of income (via changing hours) and the different gains from making such an adjustment across different kinks,

stagnant economy. This implication indicates the potential of the bunching approach in capturing the desired underlying behavioral response in some special cases, which could be explored beyond the personal income tax setting in the future.

4.5.3 Other testable implications

"Temporary bunchers" in a growing economy. In general it is natural to assume a positive wage growth rate $g > 0$ for a given individual, which is normally true for most countries, and particularly true for China. This implies that the stable bunching pattern in each period reflects temporary adjustments of hours by different groups of people. Driven by normal wage growth, people would only make a temporary income adjustment when their income happen to be falling into the bunching region. As wage rate grows steadily, or due to the fluctuations of income over time, people will not stay in the bunching region for a long time. This "temporary bunchers" implication is supported by our data. To examine this, we show the distribution of individuals by bunching months. Table 5 shows that, for each bunching kink, most bunchers stay around the kink for no more than 3 months, suggesting bunching mainly as a symptom of a temporary income adjustment, consistent with the "temporary bunchers" implication.[52] For a given group of people, bunching does not persist due to natural income growth and the fixed kink position.[53] The bunching ETI well captures the temporary income adjustment distorted by the local MTR change in a stable way. Fundamentally, it is because different people make similar responses when their incomes are around kinks.

Bimodal distribution of income adjustment. Our model assumes that people occasionally face the chance to permanently adjust hours of work, and when they do, they adjust hours permanently to their new long-run optimum. This implies that the distribution of income growth across periods should exhibit a bimodal pattern: for initial periods after the tax reform, a small portion of people facing a chance to permanently adjust hours will exhibit a larger income adjustment, while the rest of people not facing such a chance would follow their previous income path and exhibit a more stable income growth. To examine this, we focus on people experiencing MTR changes due to the tax reform. Figure 9 shows the distribution of individual average income growth from the 3-month period before the tax reform

[52] Here we use the decimal sample. Using full sample yields similar results.

[53] Even if kinks are inflation indexed as in some countries, it is unlikely that a group of people will stay around a kink for a long time, given that income growth rates of different people are quite heterogeneous and these income growth rates are probably different from the general inflation rate or the average income growth rate of all people.

Table 5: Distribution of individuals by bunching months

	Number of individuals		
Number of bunching months	Pre-reform kink 20000 RMB: 2009.6-2011.8 (27 months)	Post-reform kink 9000 RMB: 2011.9-2013.12 (28 months)	Post-reform kink 35000 RMB: 2011.9-2013.12 (28 months)
1	270	3,762	572
2	21	620	81
3	3	216	34
4	2	124	25
5	0	83	10
6	3	42	10
7	0	21	6
8	1	24	1
9	3	20	11
10	2	19	5
11	3	10	2
12	1	17	4
>12	8	28	11
Total	317	4,986	772

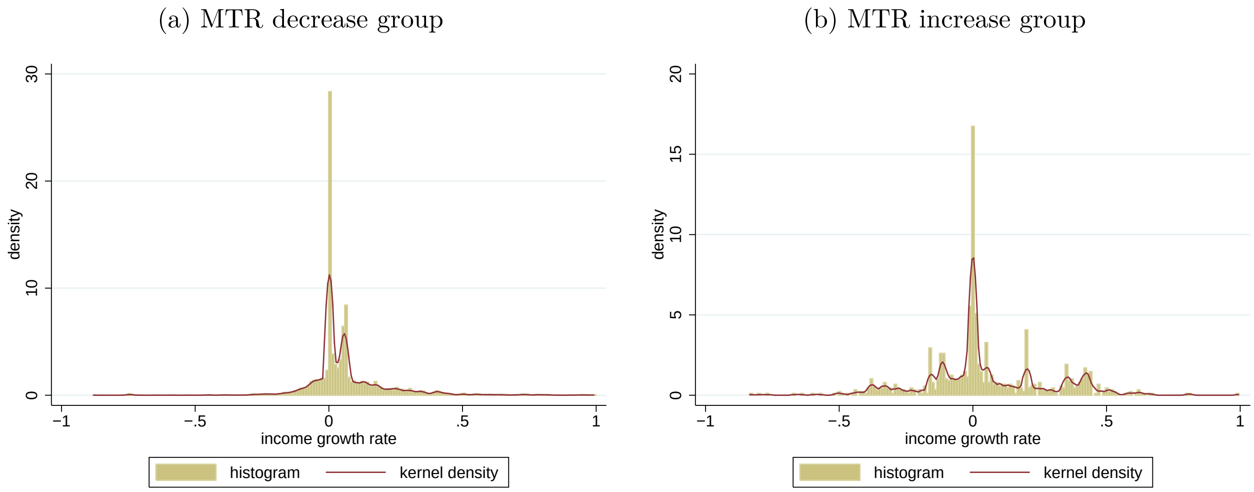
to the 3-month period after the reform. For people experiencing MTR decreases, who consist the majority of our sample, there is a clear bimodal distribution of income adjustment, thus providing support to our model. Yet for people experiencing MTR increases, the bimodal pattern is less salient, suggsting asymmetric responses to MTR increases versus decreases.

Heterogeneity. The heterogeneity among people may also have implications for bunching. In particular, people with a smaller c or ϕ would be more likely to be in the bunching region because they would face larger net gains from making the temporary income adjustment. Evidence from data provides support for this prediction. As table A5 shows, people in the bunching region tend to have a higher position than the nearby non-bunchers. These people are more likely to have managing power in the workplace (with position of middle deputy and above) and thus are facing a lower c or ϕ.

4.5.4 Alternative hypothesis

In the above section, we propose a simple model that can account for the stylized facts using parsimonious elements and generate testable predictions. In spite of this, there might be other alternative explanation that can account for the differential behavioral responses captured by

Figure 9: Bimodal distribution of income adjustment



Notes: The figure shows the distribution of individual average income growth rate from the 3-month period before the tax reform to the 3-month period after the reform. Bin width=0.01.

the two empirical approaches. One hypothesis is that while the tax reform approach captures the behavior of all people, the bunching approach mostly captures behavioral responses of those that are more able to adjust taxable income to the kink place. As we show in table A5, the bunchers at middle-high taxable income kinks tend to have a higher position than the non-bunchers, i.e. bunchers are more likely to own some managing power in the workplace (with position of middle deputy and above). This evidence does show that people who are able to bunch are not the same with those who are not. However, this hypothesis alone cannot explain either of the stylized facts. It instead provides some complementary explanation to our model. Our simple model shows how the different ETI estimates obtained by two approaches can be generated from simple assumptions even among homogeneous individuals. Given our model, admitting that bunchers are somewhat different from general persons implies that while the tax reform approach can capture permanent behavioral responses of the general people, the bunching approach could more of capture the temporary behavioral responses of individuals with a higher position, who are more able to make such adjustment.

5 Welfare implications

5.1 Deadweight loss of China's wage income tax

Total deadweight loss. We follow the formula in Feldstein (1999) to calculate the deadweight loss (DWL) of the personal wage income tax:[54]

$$DWL = 0.5\frac{\tau^2}{1-\tau}e \cdot TI.$$

Here e is the estimated ETI. In practice, DWL is calculated as

$$DWL = 0.5e\sum_i\sum_t \frac{\tau_{it}^2}{1-\tau_{it}} \cdot TI_{it}.$$

Then we can obtain the DWL as a percentage of the total tax revenue $\frac{DWL}{TR}$, which measures how much money metric welfare loss is caused by per unit tax revenue levied. Note that the DWL is proportional to e. Since our main results imply that the ETI estimates obtained from the tax reform approach are relevant for policy making, we use them for welfare calculations. Consider two cases, $e = 2.423$, the ETI estimated using the standard tax reform approach, and $e = 4$, the long-run ETI suggested by the revised tax reform approach. As table 6 shows, under current tax schedule (use 2013 data), when $e = 2.423$, $\frac{DWL}{TR}$ is 61%, which measures the 1-year average efficiency cost, when $e = 4$, $\frac{DWL}{TR}$ is 101%, which measures the long-run efficiency cost. Overall, this suggests a large efficiency loss of the current personal wage income tax. As a comparison, Feldstein (1999) calculates $\frac{DWL}{TR}$=32.2% for US personal income tax system in 1994, where his ETI estimate is 1.04.

Table 6: Deadweight loss in 2013

	current tax schedule		a 10% decrease in MTR for all people	
	e=2.423	e=4	e=2.423	e=4
DWL (million RMB)	112.38	185.52	96.27	168.72
TR (million RMB)	184.22	184.22	185.59	198.48
DWL/TR	0.61	1.01	0.52	0.85

[54]When there are transfer costs associated with sheltering (tax evasion or avoidance), then elasticity of earnings and resource cost of sheltering are needed to estimate the deadweight loss (Chetty (2009), Gorodnichenko et al. (2009)). Yet we do not have enough data to estimate these parameters. Therefore, we simply follow the standard formula by Feldstein (1999) to form a comparison with his estimates.

A 10% decrease in all marginal tax rates. Given such high ETI estimates, is it possible to improve welfare by reducing the MTR so as to reduce the deadweight loss and also increase the tax revenue as the logic of the Laffer curve implies?[55] To explore this, we take an exercise by assuming an MTR decrease by 10% for all people, following the spirit of Feldstein (1999). As table 6 shows, if $e = 2.423$, a 10% decrease in all MTRs would induce the total tax revenue to slightly increase by 0.7% ((185.59-184.22)/184.22), and the deadweight loss drops by 14.3% ((96.27-112.38)/112.38). In the long run ($e = 4$), a 10% decrease in MTRs would result in a greater revenue gain of 7.7% ((198.48-184.22)/184.22) plus a smaller decrease in the DWL of 9.1% ((168.72-185.52)/185.52). Thus, an important implication from our ETI estimates is that a further MTR decrease would be unambiguously desirable from both revenue and efficiency concerns, given the large ETI estimates in China.

5.2 Efficiency effect of the 2011 tax reform

The large magnitude of ETI estimates obtained in China suggests large efficiency cost of current personal income tax system. But what is the overall efficiency effect caused by the 2011 PIT reform? While some people experience MTR decreases, some others experience MTR increases due to the tax reform. Therefore, the overall efficiency effect induced by the tax reform is not obvious. Fortunately, it suffices to examine the revenue change due to behavioral responses in order to measure the overall efficiency effect.[56] The total revenue change due to the tax reform can be decomposed into two parts, i.e. $\triangle TR = \triangle TR|_{mechanical} + \triangle TR|_{behavioral}$, where $\triangle TR|_{mechanical}$ denotes the mechanical tax revenue change due to changes in tax code (i.e. changes in standard deduction and tax rates) while holding individual behavior unchanged, and $\triangle TR|_{behavioral}$ denotes the tax revenue change due to all potential behavioral responses, such as responses in labor supply, tax evasion, or tax avoidance. Figure 10 illustrates how we estimate $\triangle TR|_{behavioral}$. Panel (a) shows the evolution of monthly wage income tax revenue TR_t. The change in TR_t around the tax reform measures the total revenue effect $\triangle TR$. TR_t^{pred} denotes the predicted monthly total tax revenue assuming all observations were subject to the post-reform tax schedule. Before the tax reform, TR_t^{pred} were below TR_t due to a higher post-reform standard deduction and adjusted marginal tax rates. After the reform, the two figures simply coincide.[57] Then the difference between TR_t^{pred} and TR_t before the tax reform measures the mechanical revenue effect $\triangle TR|_{mechanical}$. Therefore, the change in TR_t^{pred} around the reform measures the revenue effect due to behavioral response

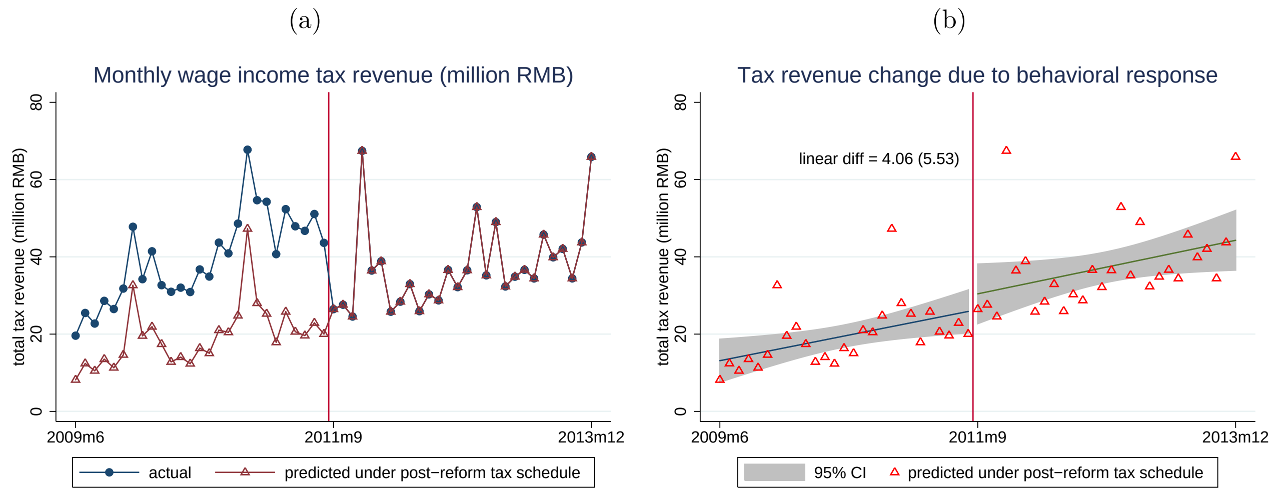
[55] Note that since we do not have a linear tax schedule, the well-known result of revenue maximizing elasticity of taxable income w.r.t. the MTR equals 1 does not hold.

[56] See Auerbach (1985), Slemrod (1998), or equation (33) of Chetty (2009).

[57] The seasonal spikes in Decembers are possibly due to bonuses release at the end of a year.

$\triangle TR|_{behavioral}$.

Figure 10: Change in tax revenue around 2011 tax reform



To estimate the efficiency effect, we use the following specification:

$$TR_t^{pred} = \alpha + \mathbf{1}[t \geq c] \cdot (g_l(t-c) + \lambda) + \mathbf{1}[t < c] \cdot g_r(c-t) + \varepsilon_t, \tag{7}$$

where c indicates the tax reform time point (September, 2011) (e.g. for October, 2011, $t-c$=1), g_l and g_r are polynomial functions in left and right hand sides of the tax reform time, α measures the left limit level of y_t around the cutoff c, λ measures the effect of the reform on the outcome variable. In practice, we adopt a linear polynomial since it fits the data pattern well. But using a higher degree polynomial does not affect our conclusion. Panel (b) of figure 10 shows that the overall efficiency effect is a positive 4.06 million RMB, though statistically insignificant (with a s.d. of 5.53 million RMB and a p-value of 0.47). This implies that despite large marginal efficiency cost as implied by the ETI estimates, the overall efficiency effect of the 2011 tax reform is close to zero, suggesting the efficiency gains from those experiencing MTR decreases (the large amount of lower earners) largely offset the efficiency losses from those experiencing MTR increases (relatively few high earners). Although redistribution (reducing tax for the poor and increasing tax for the rich) is likely be the main concern for the 2011 tax reform design, the tax reform, somewhat unexpectly, reached an effect of no aggregate efficiency effect. Usually a tax reform is designed to be revenue neutral to keep fiscal balance. But China's 2011 personal income tax reform seems to have an efficiency-neutral property at the cost of an immediate revenue loss (clear from panel (a) of figure 10).

6 Conclusion

Bunching is a recently developed approach that has great potential to be applied in many fields such as taxation, social security, social insurance, welfare programs, and others.[58] While the techniques to estimate desired parameters from observed bunching are quite mature, the underlying behavioral responses captured by bunching are not all that clear. In this paper, we show that under personal income tax setting, the bunching approach captures very different behavioral response to marginal tax rate changes than that captured by the tax reform approach. We document that while the tax reform ETI estimates increase concavely over time, the bunching ETI estimates are stable and much smaller. While the tax reform approach can capture the infrequent but permanent adjustment of hours of work to tax changes, the bunching approach can only reflect temporary adjustment. We show that a simple model reconciles the sharp contrasts between ETI estimates obtained by the two approaches.

Of course, beyond the personal income tax context, bunching does not necessarily capture only temporary behavioral adjustments. In other cases, bunching may be able to reflect the longer-run behavioral responses. In fact, our model implies that in a stagnant economy, the bunching ETI can capture the permanent adjustment of behavior in the long run. But in most cases, bunching estimates alone are not sufficient to reveal the underlying structural parameter, as the adjustment costs that differentiate the bunching estimates from the tax reform estimates are not able to be pinned down solely by the bunching ETI. A lesson from this paper is that the behavioral responses underlying bunching should be carefully examined case by case and a policy implication from the bunching estimates should be cautiously made.

Our findings show the limitation of bunching ETI estimates for policy concerns and therefore recommend the tax reform ETI as more relevent for policy making. In addition, the large ETI estimates obtained in this paper suggest that marginal tax rates are too high in China. We show that a uniform MTR decrease would bring about both revenue increase and efficiency improvement, at least for our sample city. This clearly suggests reducing the MTRs in a future personal income tax reform. Any argument to increase the MTR for high income earners for redistributional reason does not receive support from our ETI estimates.

[58]Kleven (2016) provides an up-to-date review on potential applications of the bunching approach.

References

Auerbach, A. J.: 1985, 'The Theory of Excess Burden and Optimal Taxation'. *Handbook of Public Economics* **1**, 61–127.

Bastani, S. and H. Selin: 2014, 'Bunching and non-bunching at kink points of the Swedish tax schedule'. *Journal of Public Economics* **109**, 36–49.

Benzarti, Y., D. Carloni, J. Harju, and T. Kosonen: 2017, 'What Goes Up May Not Come Down: Asymmetric Incidence of Value-Added Taxes'.

Best, M. C. and H. J. Kleven: 2016, 'Housing Market Responses to Transaction Taxes: Evidence From Notches and Stimulus in the UK'.

Blomquist, S. and W. K. Newey: 2017, 'The Bunching Estimator Cannot Identify the Taxable Income Elasticity'.

Blomquist, S. and H. Selin: 2010, 'Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates'. *Journal of Public Economics* **94**(11), 878–889.

Chetty, R.: 2009, 'Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance'. *American Economic Journal: Economic Policy* pp. 31–52.

Chetty, R.: 2012, 'Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply'. *Econometrica* **80**(3), 969–1018.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri: 2011, 'Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records'. *The Quarterly Journal of Economics* **126**(2), 749–804.

Chetty, R., J. N. Friedman, T. Olsen, L. Pistaferri, et al.: 2010, 'Adjustment costs, firm responses, and labor supply elasticities: evidence from Danish tax records'.

Devereux, M. P., L. Liu, and S. Loretz: 2014, 'The Elasticity of Corporate Taxable Income: New Evidence from UK Tax Records'. *American Economic Journal: Economic Policy* **6**(2), 19–53.

Fack, G. and C. Landais: 2016, 'The effect of tax enforcement on tax elasticities: Evidence from charitable contributions in France'. *Journal of Public Economics* **133**, 23–40.

Feldstein, M.: 1995, 'The effect of marginal tax rates on taxable income: a panel study of the 1986 Tax Reform Act'. *Journal of Political Economy* pp. 551–572.

Feldstein, M.: 1999, 'Tax avoidance and the deadweight loss of the income tax'. *Review of Economics and Statistics* **81**(4), 674–680.

Gelber, A. M., D. Jones, and D. W. Sacks: 2015, 'Earnings adjustment frictions: Evidence from the Social Security earnings test'.

Gordon, R. H. and J. B. Slemrod: 2000, 'Are "Real" Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases?'. *Does Atlas Shrug?: The Economic Consequences of Taxing the Rich* pp. 240–280.

Gorodnichenko, Y., J. Martinez-Vazquez, and K. S. Peter: 2009, 'Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia'. *Journal of Political Economy* **117**(3), 504–554.

Gruber, J. and E. Saez: 2002, 'The Elasticity of Taxable Income: Evidence and Implications'. *Journal of Public Economics* **84**(1), 1–32.

Harju, J. and T. Matikka: 2016, 'The Elasticity of Taxable Income and Income-Shifting: What is "Real" and What is Not?'. *International Tax and Public Finance* pp. 1–30.

Ito, K.: 2014, 'Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing'. *The American Economic Review* **104**(2), 537–563.

Jantti, M., J. Pirttila, and H. Selin: 2015, 'Estimating labour supply elasticities based on cross-country micro data: A bridge between micro and macro estimates?'. *Journal of Public Economics* **127**, 87 – 99. The Nordic Model.

Kleven, H. J.: 2016, 'Bunching'. *Annual Review of Economics* **8**, 435–464.

Kleven, H. J. and E. A. Schultz: 2014, 'Estimating taxable income responses using Danish tax reforms'. *American Economic Journal: Economic Policy* **6**(4), 271–301.

Kleven, H. J. and M. Waseem: 2013, 'Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan'. *The Quarterly Journal of Economics* **128**(2), 669–723.

Kopczuk, W.: 2005, 'Tax bases, tax rates and the elasticity of reported income'. *Journal of Public Economics* **89**(11), 2093–2119.

Kreiner, C. T., S. Leth-Petersen, and P. E. Skov: 2016, 'Tax Reforms and Intertemporal Shifting of Wage Income: Evidence from Danish Monthly Payroll Records'. *American Economic Journal: Economic Policy* **8**(3), 233–57.

Le Maire, D. and B. Schjerning: 2013, 'Tax bunching, income shifting and self-employment'. *Journal of Public Economics* **107**, 1–18.

Peterman, W. B.: 2016, 'Reconciling micro and macro estimates of the Frisch labor supply elasticity'. *Economic Inquiry* **54**(1), 100–120.

Saez, E.: 2003, 'The Effect of Marginal Tax Rates on Income: A Panel Study of 'Bracket Creep''. *Journal of Public Economics* **87**, 1231–1258.

Saez, E.: 2010, 'Do taxpayers bunch at kink points?'. *American Economic Journal: Economic Policy* pp. 180–212.

Saez, E., J. Slemrod, and S. H. Giertz: 2012, 'The Elasticity of Taxable Income With Respect to Marginal Tax Rates: A Critical Review'. *Journal of Economic Literature* pp. 3–50.

Saez, E., J. B. Slemrod, and S. H. Giertz: 2009, 'The Elasticity of Taxable Income With Respect to Marginal Tax Rates: A Critical Review'. *National Bureau of Economic Research Working Paper 15012*.

Slemrod, J.: 1998, 'Methodological issues in measuring and interpreting taxable income elasticities'. *National Tax Journal* pp. 773–788.

Slemrod, J. and S. Yitzhaki: 2002, 'Tax Avoidance, Evasion, and Administration'. *Handbook of Public Economics* **3**, 1423–1470.

Weber, C. E.: 2014, 'Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences'. *Journal of Public Economics*.

Online Appendix

(Not For Publication)

A China's personal income tax schedule

There are 11 income components taxed under China's personal income tax. They are divided into three broader components: wages/salaries income, self-employment income, and other incomes. Self-employment incomes include (1) income from production or business operation by self-employed industrial and commercial households, and (2) income from the contract and operation of enterprises and institutions and the business income of the lease. Other income components include (1) remuneration for providing services,[59] (2) author's remuneration, (3) royalties income, (4) interest, dividend, and bonuses, (5) income from leasing property, (6) property transfer income, (7) accidental income, and (8) other incomes. Table A1, figure 1, and figure A1 show the tax rates, standard deductions, and changes in them due to the 2011 PIT reform. Overall, the 2011 tax reform increases the monthly standard deduction for monthly wage/salary from 2,000 RMB to 3,500 RMB, increases the MTR for high wage/salary earners and decrease the MTR for low wage/salary earners. For the self-employment incomes, the statutory MTR generally decreases for all income levels that experience an MTR change. But since most self-employed businesses pay a pre-determined fixed amount income tax, it is not clear how their tax liabilities changed in reality.

Self-employment income tax. Despite the statutory progressive tax rate, as shown in figure A1, in practice, based on the availability of a reliable accounting book, the rule to tax self-employment income is as follows. First, if a self-employed business has a complete accounting book on costs and incomes and the tax bureau finds it reliable, then self-employment income implied by the accounting book will be taxed following the progressive tax rate structure. This approach, called *chazhang* levy, is usually applied to large scale self-employed businesses and is very rarely used in practice. Second, for most self-employed businesses, since the accounting book is largely absent or severely incomplete, it is hard to obtain a precise measure of self-employment income. Accordingly, the tax bureau adopt the so-called

[59] Remuneration for providing services is paid to part-time work and enjoy a standard deduction of 800 RMB if this income does not exceed 4,000 RMB (see table A1), while wage/salary is paid to full-time work. Suppose a person has a wage earnings of 2,800 RMB before the tax reform, he can deduct 2,000 RMB and his taxable income is 800 RMB. If he instead has a wage earnings of 2,000 RMB plus a remuneration for providing services of 800 RMB, then he can deduct 2,000 RMB for wages and deduct 800 RMB for remuneration for providing services, which results in a zero taxable income. There is anecdotal evidence that some people may use the deduction by income item to avoid tax.

heding levy approach, i.e. tax levy based on assessment. This levy approach first requires an assessment of business volume based on business area, industry, measurable business costs (e.g. electricity and water), receipts (if any), and other useful information. Then based on the assessment, there are two tax levy approaches. The first approach is to generate taxable income based on the assessed income and a taxable income ratio varying by industry, and then obtain the tax liability by imposing the progressive tax rate on this taxable income. The second approach, which is applied to the majority of self-employed businesses, is a predetermined fixed amount tax approach, called *dinge* levy.[60] In particular, the tax bureau will make a list of all self-employed businesses that are subject to this approach and the fixed amount of personal income tax liability for each of them.[61] [62] In many cities, this list is required to be made publicly available online periodically. Theoretically, this predetermined tax rule might cause some distortion in firm's behavior, though possibly much less than that would have been caused by a progressive tax rate structure.[63] Given this background, and due to the lack of self-employment income data, throughout this paper, we mainly focus on wage/salary income. This is approapriate since our main focus is the ETI w.r.t. tax rate and the self-employment income tax is not based on a rate structure. In addition, for the majority of people, wage/salary income is the major income source.

[60]We do not find figures on ratios of self-employed business subject to different tax levy approach. By by inquiring tax bureau staff, we confirm that for our sample city and all the other cities we have asked, most self-employment income is taxed under the *dinge* levy approach. The wide adoption of this approach may be due to its relatively low enforcement cost as compared to its alternatives.

[61]Each local tax bureau has its own way to determine the amount of this predetermined fixed amount tax. The formula is not made public. But it is said that the fixed amount may depend on a self-employment business's previous income, industry, and other useful information.

[62]A large fraction of self-employed businesses have a too small scale to pay any personal income tax. For example, currently in Anhui province, if the assessed monthly business revenue is less than 30,000 RMB, then no personal income tax needs to be paid. By inquiring relevant tax stuff, we know that in Jiang county of Shanxi province, only 10% of self-employed businesses need to pay personal income tax.

[63]In China, in addition to the personal income tax, self-employed businesses are also subject to many other taxes and fees including the value-added tax, business tax, city maintenance and construction tax, local education supplementary fee, stamp tax, housing property tax, operations tax, etc. In practice, all these taxes for the self-employed follow the *dinge* levy approach, i.e. having a predetermined fixed amount property.

B Tax on bonuses

In China, bonuses are taxed separately and differently from regular monthly wage/salary. In the following, we first show how bonuses are taxed differently from regular wage, and then explore the implications of such difference.

Regular monthly wage follows the tax schedule in figure 1, which can also be summarized by table A3. In practice, tax liability is calculated using table A3 for its convenience. For example, for post-reform period, after obtaining the taxable income of a month, the tax liability is calculated as $TI \times MTR$ - quick deduction number.

Bonuses are not taxed in the same way as regular wage, since otherwise the tax law would impose a penalty on bonuses, given bonuses are normally much higher than regular wage and the tax rate structure is graduated. In the month when the bonuses are paid, regular wage and bonuses are taxed separately. Suppose bonuses in a month are B RMB, the regular wage in this month is W RMB, the standard deduction is D RMB. There are two cases. (1) if W>D, then W is taxed in regular way. For B, we first divide it by 12, and then find the corresponding MTR in table A3, the tax liability for B is then ($MTR \times B$ - quick deduction number) RMB. (2) if W<D, then W is not taxable. For B, we first divide B+W-D by 12, then find the corresponding MTR in table A3, the tax liability is then ($MTR \times (B + W - D)$ - quick deduction number) RMB.

By dividing bonuses by 12 before applying the MTR of the regular wage tax schedule, this "tax on bonuses" rule is designed to avoid the unusually high MTR pushed up by high bonuses. However, the formula of calculating tax liability for bonuses is inappropriately designed in the sense that it creates tax notches (i.e. there is a jump in total tax liability when income is slightly higher than the cutoff value). The problem lies in the inappropriate application of quick deduction number in regular wage tax formula to the bonuses tax formula. For example, in the post-reform period, if bonus is slightly lower than 18,000 RMB, then an MTR of 3% applies, and the tax liability is slightly lower than $18,000 \times 3$ RMB. But if bonus is slightly higher than 18,000 RMB, then an MTR of 10% applies, and the tax liability is slightly higher than $18,000 \times 10$ RMB. Interestingly, although there is evidence that the tax notches for bonuses are well-known by shrewd accounting staffs (we see clear bunching of bonuses at these notches), the tax formula for bonuses has never been corrected in tax law.

Our data do not indicate which income corresponds to bonuses. But we are able to observe which incomes apply the "tax on bonuses" rule. We thus define the incomes following the "tax on bonuses" rule as bonuses.

If bonuses are taxed in the same way as regular monthly wage, then the optimal wage and bonuses distribution (in the sense of paying least tax), for a given annual earnings, would pay largely equal wage and bonuses across months. Given the "tax on bonuses" rule, what is the combination of wage (W) and bonuses (B) that yields least tax, for a given annual earnings (E)? Let's take post-reform period (e.g. 2012 and 2013) for example, which corresponds to a standard deduction of 3,500 RMB. If E is less than $3,500 \times 12 = 42,000$ RMB, a combination of $W = E$ and $B = 0$ yield tax liability 0. If E lies between 42,000 RMB and 78,000 RMB ($= 42,000 + 1,500 \times 12 + 18,000$), then the optimal distribution is some combination between wage and bonuses so that the MTR for bonuses and regular wages are no more than 3%. The tax liability under the optimal distribution is $(E - 42,000) \times 3\%$ RMB, though there are infinite possible combinations of W and B to achieve this goal. For optimal bonuses policy and lowest tax liability for a given annual earnings, we solve the following tax miminization problem:

$$min_{W,B}\, T_1(W) + T_2(B)\, s.t.\, W + B = E,$$

where

$$T_1(W) = \begin{cases} 0 & if\, W \in [0, 42000] \\ (W - 42000) \times 0.03 & if\, W \in (42000, 60000] \\ (W - 42000) \times 0.1 - 12 \times 105 & if\, W \in (60000, 96000] \\ (W - 42000) \times 0.2 - 12 \times 555 & if\, W \in (96000, 150000] \\ (W - 42000) \times 0.25 - 12 \times 1005 & if\, W \in (150000, 462000] \\ (W - 42000) \times 0.3 - 12 \times 2755 & if\, W \in (462000, 702000] \\ (W - 42000) \times 0.35 - 12 \times 5505 & if\, W \in (702000, 1002000] \\ (W - 42000) \times 0.45 - 12 \times 13505 & if\, W \in (1002000, \infty), \end{cases}$$

and

$$T_2(B) = \begin{cases} B \times 0.03 & if\, B \in [0, 18000] \\ B \times 0.1 - 105 & if\, B \in (18000, 54000] \\ B \times 0.2 - 555 & if\, B \in (54000, 108000] \\ B \times 0.25 - 1005 & if\, B \in (108000, 420000] \\ B \times 0.3 - 2755 & if\, B \in (420000, 660000] \\ B \times 0.35 - 5505 & if\, B \in (660000, 960000] \\ B \times 0.45 - 13505 & if\, B \in (960000, \infty). \end{cases}$$

Solving this problem, we can obtain the lowest tax liability as

$$T(E) = \begin{cases} 0 & if\ E \in [0, 42000] \\ 0.03E - 1260 & if\ E \in [42000, 78000] \\ 0.1E - 6720 & if\ E \in (78000, 114000] \\ 0.2E - 18120 & if\ E \in (114000, 125550] \\ 0.1E - 5565 & if\ E \in (125550, 150000] \\ 0.2E - 20565 & if\ E \in (150000, 204000] \\ 0.25E - 30765 & if\ E \in (204000, 516000] \\ 0.3E - 56565 & if\ E \in (516000, 565500] \\ 0.2E - 15 & if\ E \in (565500, 570000] \\ 0.3E - 57015 & if\ E \in (570000, 669000] \\ 0.25E - 23565 & if\ E \in (669000, 862000] \\ 0.3E - 67665 & if\ E \in (862000, 1122000] \\ 0.35E - 123765 & if\ E \in (1122000, 1422000] \\ 0.45E - 265965 & if\ E \in (1422000, 1550333] \\ 0.3E - 33415 & if\ E \in (1550333, 1662000] \\ 0.45E - 282715 & if\ E \in (1662000, \infty), \end{cases}$$

and the optimal bonuses policy is

$$B(E) = \begin{cases} 0 & if\ E \in [0, 42000] \\ [0, E - 42000] & if\ E \in (42000, 60000] \\ [E - 60000, 18000] & if\ E \in (60000, 78000] \\ 18000 & if\ E \in (78000, 125550] \\ E - 96000 & if\ E \in (125550, 150000] \\ 54000 & if\ E \in (150000, 565500] \\ E - 462000 & if\ E \in (565500, 570000] \\ 108000 & if\ E \in (570000, 669000] \\ E - 462000 & if\ E \in (669000, 862000] \\ 420000 & if\ E \in (862000, 1550333] \\ E - 1002000 & if\ E \in (1550333, 1662000] \\ 660000 & if\ E \in (1662000, \infty). \end{cases} \quad (8)$$

The lowest tax liability $T(E)$ is a continuous function of E. But with tax notches for bonuses, the optimal bonuses policy is not a continuous function of annual earnings E. From equation 8, we expect to see significant bunching of bonuses at values like 0 RMB, 18,000 RMB, 54,000 RMB etc. For annual earnings between 42,000 RMB and 78,000 RMB, there are infinite possible optimal bonuses falling in a certain range, though the optimal bonuses policy is unique for the other annual earnings. Would actual bonuses be consistent with equation 8? First, we find surprisingly few incomes applying the "tax on bonuses" rule. In 2013, only 0.45% (2,492 in 550,506 individuals) of people have any bonuses in our data.[64] This fraction is very similar in other years. Equation 8 shows that theoretically people with annual earnings above 42,000 RMB should have positive bonuses. In 2013, there are 192,893 individuals having annual earnings above 42,000 RMB. Hence the few people having bonuses cannot be explained by the optimal bonuses policy. We are not entirely sure why there are so few people having bonuses. Perhaps many people receive bonuses in cash, as much anecdotal evidence says. Now let us just focus on those having positive bonuses, and see if they follow the optimal bonuses policy.

In figure A10, we compare the optimal bonuses policy and actual bonuses in 2013. We focus on annual earnings below the threshold of 565,500 RMB because almost all individuals with positive bonuses in 2013 have annual earnings below this threshold. Figure A10 shows that although actual bonuses do not coincide the optimal bonuses policy in a precise manner, they follow similar pattern. In particular, like the optimal bonuses policy, for a wide range of annual earnings starting from 0 RMB, the actual bonuses remain at a largely constantly low level, corresponding to the optimal bonus policy of 0 for initial levels of earnings. Then at the middle range of incomes, there is a linear relation between bonuses and annual earnings, similar to the optimal bonuses policy. Finally, when annual earnings surpasses certain level, the bonuses is around 54,000 RMB, which is exactly predicted by the tax minimization problem. Overall, the actual distribution of bonuses is a right-shift of the theoretical prediction. This is reasonable because the theoretical prediction is made by assuming there is only standard deduction, while in fact, the existence of other deductions and tax-exempt incomes would make the actual distribution shift right.

Of course, there are realistic concerns that actual bonuses may not follow the optimal distribution predicted in equation 8. For one, obviously, the tax minimization problem is quite complicated to solve. It might be hard to imagine that employers, even with the help of accounting staffs, make such complicated calculations when deciding how much wages and bonuses to pay. For another, in reality, people (even employers) may not be able to fully anticipate the annual earnings for each employee and make optimal arragement from the

[64]We restrict our sample to people with positive income in every month in 2013.

very beginning. It is possible that bonuses are a function of random shock that cannot be fully anticipated when the first several monthly wages are paid. Due to these concerns, it is possible to see the actual tax liability and amount of bonuses differ from the optimal one, for a given annual earnings.

Figure A10 does show some evidence for imperfect optimization, consistent with the above concerns. First, we do see non-zero bonuses paid for bottom levels of incomes, consistent with the hypothesis that annual earnings is not perfectly anticipated at the very beginning. Second, the actual distribution of bonuses does not follow the discrete distribution as predicted by equation 8. In particular, it seems to "smooth out" the highly nonlinear part in the optimal distribution (corresponding to earnings between 42,000 RMB and 125,550 RMB). There is no evidence for a flat part of bonuses at 18,000 RMB, as predicted by equation 8. But there is clear evidence for a flat part of bonuses at 54,000 RMB. How to reconcile these? The lack of evidence for the part of non-linear distribution of bonuses is probably due to the inability of making perfect solution of the tax minimization problem. In addition, the income range corresponding to the optimal bonuses at 18,000 RMB is much narrower than that corresponding to 54,000 RMB, suggesting the benefit of making correct optimization is much smaller in the former case than in the latter case. This explains the lack of a flat part of bonuses at 18,000 RMB and the existence of a flat part of bonuses at 54,000 RMB.

To sum up, in this section, we explore the theoretically optimal bonuses policy for a given annual earnings, and find that the actual distribution of bonuses follows the theoretical distribution in a smoother way. The places where the actual distribution of bonuses differ from the theoretically optimal distribution imply an imperfect optimization mode, consistent with the complicated and seperate tax schedules for bonuses and regular wage. However, our data seem to have too few incomes taxed as bonuses to make it an important concern for our main analysis. Thus, we do not account for the potential complexity brought about by bonuses in the main text.

C Bunching evidence of full sample and decimal sample

Bunching evidence for full sample. Figure A3 displays evidence of bunching for full sample (including both round and decimal values) of taxable wage/salary income. There is clear evidence of excess bunching at kinks, in both pre-reform sample (2009.6-2011.8) and post-reform sample (2011.9-2013.12). There is also clear bunching in some non-kink points (e.g. 1,000, 4,000, 6,000 in pre-reform sample and 1,500, 3,500 in post-reform sample). This is because employers may have a tendency to report taxable incomes at "rounder" numbers ending in hundreds, thousands etc. Bunching at these places create a problem for us to reveal the "true" bunching scale caused by kinks. In previous literature, usually observed is a consistently smaller regular bunching pattern at non-kink rounder numbers (Kleven and Waseem (2013) Devereux et al. (2014) Best and Kleven (2016)). Then the relative bunching scale at kinks compared to that at non-kink rounder numbers would nicely reveal the bunching ETI estimate. But in our data, bunching patterns at non-kink rounder numbers are so irregular that we cannot use them as a reference to reveal the bunching scale caused by kinks. For example, for the pre-reform kink at 5,000 RMB, if we (reasonably) use bunching at 4,000 RMB and 6,000 RMB as reference to predict the counterfactual income distribution at 5,000 RMB, then it implies a close to zero ETI at kink 5,000 RMB. But one might argue why not use bunching at 4,500 and 5,500 as reference instead. These bunching patterns in our data imply that the standard bunching approach (even incorporating the "rounder"-number bunching as did by Kleven and Waseem (2013)) does not promise an ideal way to obtain the counterfactual taxable income distribution and thus cannot generate an ideal estimation of ETI using full sample.

In addition, the full sample bunching evidence also shows some unusual excess bunching points (e.g. taxable incomes of 3,637.5 RMB, 8,170 RMB, 9,069 RMB in the post-reform sample) that are not multiples of the salient round numbers (e.g. 50, 100, 500, 5000). Bunching at these taxable incomes do not receive any reasonable theoretical support. A careful examination of these unusual bunching points reveals that these unusual bunching points probably result from a one-shot wage release by certain employers.[65] In our bunching analysis, we need to exclude these outliers.

To address the above problems, we are motivated to seek for evidence of bunching at places that are systematically unaffected by such problem. Fortunately, when we focus on the decimal sample (i.e. those with decimal taxable incomes), these irregular bunching patterns

[65]For example, almost all taxable incomes of 3,637.5 (8,170) RMB in the post-reform sample is associated with pre-tax wage income of 7,137.5 (12,070) RMB, all of which are reported in May, 2012 (December, 2011).

are exlcuded.

Bunching evidence for decimal sample. Figure A4 shows bunching evidence for the decimal sample. It is clear that during pre-reform period, there is no bunching at all in kinks of 500 RMB, 2,000 RMB, and 5,000 RMB. There is clear bunching at 20,000 RMB. Bunching at higher kinks is also clear, although observations are too few to render precise estimates. For the post-reform sample, there is slight bunching evidence at 1,500 RMB and 4,500 RMB. Bunching is clear at 9,000 RMB and 35,000 RMB. Density at higher kinks is too noisy to render clear bunching evidence. Due to these graphical evidence, we focus on pre-reform kink at 20,000 RMB and post-reform kinks at 9,000 RMB and 35,000 RMB.

The lack of bunching at bottom TI kinks in the decimal sample is consistent with the comparable bunching scale at these kinks to neighboring non-kink places in the full sample (e.g. compare bunching at 5,000 RMB with that at 4,000 RMB and 6,000 RMB). And the bunching pattern at higher kinks in the decimal sample are similar to that in the full sample. This suggests restricting to a decimal sample would well excludes the irregular bunching at non-kink places while reveal desired bunching pattern sat kinks.

D Placebo bunching evidence: comparing bunchings before and after the tax reform

In this section, we use pre-reform and post-reform samples to examine bunching at the new kinks and old kinks. We show that there was no bunching at the newly imposed kinks before the tax reform. We also show that after an old kink was abolished, bunching disappeared within a short time.

Bunching of taxable income adjusted by standard deduction. In 2011 tax reform, not only the MTR brackets of taxable income changed, the standard deduction for wage incomes also changed. Thus, the post-reform bunching at kinks may correspond to different levels of pre-deduction wage incomes from the pre-reform ones. To account for this, we conduct a bunching analysis focusing on taxable incomes adjusted by the change of standard deduction. Note that standard deduction is 2,000 RMB before the reform and 3,500 RMB after the reform. Then for pre-reform observations, we deduct 1,500 RMB from their taxable income to render predicted post-reform taxable income. Similarly, for post-reform observations, we add 1,500 RMB to their taxable income to render predicted pre-reform taxable income. Such approach would hold other deductions and exempt incomes fixed, which were not affected by the reform, and would arguably render comparable predicted taxable income for pre-reform and post-reform samples.

We still consider pre-reform kink at 20,000 RMB and post-reform kinks at 9,000 RMB and 35,000 RMB and examine bunching at these kinks within one year before (2010.9-2011.8) and after (2011.9-2012.8) the tax reform. As figure A5 shows, there was no bunching evidence at pre-reform kink 20,000 RMB after the reform and no bunching evidence at the post-reform kinks 9,000 RMB and 35,000 RMB before the tax reform, thus providing a placebo test for the bunching at kink approach.

Bunching of unadjusted taxable income. As another placebo test, in figure A6, we examine whether bunching of *unadjusted* taxable income (1) was absent at post-reform kinks before the kinks were imposed and (2) whether it disappeared at the pre-reform kink after the kink was abolished. As expected, within one year prior to the reform, there was no evidence of bunching at 9,000 RMB (panel (c)) and 35,000 RMB (panel (d)). For the pre-reform kink 20,000 RMB (panel (a)), however, within one year after the kink was abolished, there was still clear evidence for excess bunching. This is quite unexpected because these excess bunchers should not be the same persons bunching at the same kink before the tax reform, due to the standard deduction change. Moreover, the individuals with post-reform taxable income around 20,000 RMB should have pre-reform taxable income around 21,500 RMB, assuming their raw income did not change and other exempt incomes and deductions fixed. Since 21,500 RMB is not a kink point before the tax reform, the excess bunching of taxable income at 20,000 RMB cannot be due to pre-reform bunching at kink. Is it possible that these excess bunchers were due to inattention of kink point change? If this is true, we expect to see the excess bunching persisted within a short time after the tax reform and later disappeared. However, by dividing the one year after the tax reform into two half-year periods (panel (b)), we see that within half year after the tax reform there was no bunching at the old kink, while the excess bunching emerged during the second half year after the reform. A further examination reveals that the excess bunching happened only in 2012.3. Thus we interpret it as a coincidence (possibly a firm issuing similar amount of wage income in that month) rather than systematic inattention of kink point change.

Adjustments of bunching behavior. How fast did people respond to the tax reform? To explore this, we cluster 3-month data because observations would be too scarce at these middle-high income levels if we examine each month separately.[66] Figure A7 panel (a) shows that bunching at pre-reform 20,000 RMB kink disappeared immediately after the tax reform. There is no bunching at all within 3 months after the tax reform, which suggests that the excess bunchers were well informed of the kink information both before and after the reform. While panel (b) shows fuzzy evidence on how bunching increased after 9,000 RMB taxable income became the new kink, panel (c) shows clearly that bunching increased over time at 35,000 RMB kink, suggesting it takes time to adjust wage income due to initial inattention and adjustment costs. The 2011.12-2012.2 period has more observations probably due to release of bonus at the end of a year (December).

[66] Best and Kleven (2016) in another setting use monthly figures and show sharper evidence on how bunching in housing market changes in response to stamp duty tax changes.

E Uncover structural elasticity using Chetty's bound approach

While we estimate the underlying elasticity and optimization friction using a structural model, Chetty (2012) takes an alternative approach. He bounds structural parameters without identifying how frictions affect behavior. The motivation for using a bound approach is that we cannot always ensure that the parametric specification we adopt is correct. The bound approach thus allows us to estimate the scope of the underlying elasticity with much weaker assumptions. The cost is that we in general need large tax reforms to obtain tight bounds, which is not well satisfied in our case.

We simply take his approach to estimate the bounds and compare with our main structural estimates. The closed form presentation for bounds on the structural elasticity of taxable income ε, as obtained in Chetty (2012), is $\varepsilon_L = \widehat{e} + \frac{4\delta}{(\triangle log(1-\tau))^2}(1-\rho)$ and $\varepsilon_U = \widehat{e} + \frac{4\delta}{(\triangle log(1-\tau))^2}(1+\rho)$, where $\widehat{e}$ is the observed elasticity, δ is the degree of friction, τ is marginal tax rate, $\rho = (1+\frac{1}{2}\frac{\widehat{e}}{\delta}(\triangle log(1-\tau))^2)^{1/2}$. Empirically, δ is measured by the utility cost of ignoring tax change as a fraction of net earnings. Chetty (2012) find that $\delta = 1\%$ may be a reasonable estimate for most cases and we also follow such practice. When we have multiple (say J) observed elasticities, we can also obtain the greatest lower bound $\varepsilon_L^{max} = max(\varepsilon_L^j)$ and the least upper bound $\varepsilon_U^{min} = max(\varepsilon_U^j)$.

Table A4 shows very wide bounds for the underlying elasticity. The unified bounds are $\varepsilon_L^{max} = 0.01$ and $\varepsilon_U^{min} = 17.62$. This is not surprising, since Chetty (2012) shows that with small marginal tax rate changes, in general very wide bounds for intensive margin elasticities would be obtained. Thus, in our case, the bound estimates are not very informative and we would stick to the underlying elasticity estimated using the parametric approach. Of course, the underlying elasticity estimated using the parametric approach lies in the above bounds.

F Tax reform ETI vs. bunching ETI: full sample estimates

Our main bunching ETI estimates shown in figure 6 are based on the decimal sample. There may be concern that the decimal sample could underestimate the ETI since it might exclude taxable incomes adjusted to exactly at the kink more than those with other integer values. We address this concern here by showing that using full sample would generate similar dynamic pattern of bunching estimates.

First, using full sample, we apply the revised bunching approach to estimate the global bunching ETI using bunching at kinks 9,000 RMB and 35,000 RMB, as they are the kinks used for estimating the global bunching ETI in figure 6. Since we see no bunching at these two kinks before the tax reform, it is not necessary to use the pre-reform taxable income distribution at these two kinks to construct counterfactual density distribution.[67] Figure A8 shows that the global bunching ETI estimates using full sample are very similar to those using the decimal sample.

Second, we do observe bunching increases at kinks 1,500 RMB and 4,500 RMB in post-reform compared to pre-reform periods. Since there is pre-reform bunching at these two kinks, we need to use it to construct the counterfactual density for post-reform sample.[68] For each kink, we first obtain the fraction of observations bunching at the kink in a given neighborhood around the kink in pre-reform period. Then we use this fraction to adjust the observed bunching at each post-reform period. Then we apply the standard bunching approach to obtain the an ETI for each kink in each post-reform period. Finally, using the observed ETI at all post-reform kinks 1,500 RMB, 4,500 RMB, 9,000 RMB, and 35,000 RMB (we do not estimate ETI for top kinks due to too few observations), we estimate the global bunching ETI using the above approach. Figure A8 shows that the dynamic global bunching ETI estimates are much more precisely estimated using full sample. They are smaller than the global ETI estimated only using middle-high kins. Most importantly, we again find that the dynamic global bunching ETI estimates are stable over time.

[67]It is even undesirable to do so because the monthly (or three-month pooled) taxable income distribution around these kinks are volatile across periods due to not so many observations in these middle-high kinks. Using pre-reform taxable income desity as counterfactual distribution for post-reform sample generates unreasonable estimates for these middle-high taxable income kinks.

[68]Since there are large amounts of observations at these two kinks in each period, the density distribution is rather stable across period and using pre-reform distribution to construct the post-reform counterfactual density is reasonable.

G Real responses versus income manipulations captured by tax reform and bunching ETI estimates

Traditionally, the behavioral responses to tax changes are usually divided into real responses (e.g. working hours, effort) and various income manipulations (e.g. income shifting, income underreporting). It is beyond the scope of this paper to exhaustively decompose all behavioral responses accounting for each ETI estimate. A comprehensive decomposition is also very difficult, if not impossible, as well realized in literature (e.g. Feldstein (1999)). In spite of this, revealing potential evidence on real responses versus income manipulations captured by each approach is useful because income manipulations and real responses in labor supply have different normative implications (Slemrod and Yitzhaki (2002), Chetty (2009)). It is argued that only real responses are relevant for optimal taxation, and that a good tax system should be devised to minimize the space for income manipulations. Thus if there is no evidence for real responses, then all the behavioral responses captured by the ETI estimates are simply a reporting behavior, implying small efficiency costs brought about by "real" distortions; at the same time, the large ETI estimates would also imply that the tax system has too many loops for income manipulations and should be amended correspondingly.

Tax reform ETI. Although we do not provide direct evidence of income manipulations using our data, literature contains much evidence that income manipulation can be an important factor for the tax reform ETI estimates. An extreme case is Kreiner et al. (2016), who find that the intertemporal wage income shifting can account for almost all of the ETI estimate. Income manipulations in response to tax changes could take other forms, e.g. shifting income to bases with a lower tax rate.[69] There is also anecdotal evidence, via communication with Chinese employees, that firms sometimes pay cash as part of compensation, probably out of tax evasion purpose.

To explore potential real responses to tax, we resort to survey data and examine the response of working hours. In particular, we use CFPS (China Family Panel Studies) data in 2010 and 2012 as they are the most appropriate data for our research purpose.[70] We divide individuals into three groups due to their 2010 pre-tax monthly wage income. People with wages lower than 2,000 RMB face zero tax rate before and after the tax reform. If we only consider standard deduction, for simplicity, then people with wages in 2,000 RMB-8,000 RMB interval largely correspond to those facing MTR decreases, people with monthly wages

[69] See Gordon and Slemrod (2000) for evidence of such reclassification of income in the US.

[70] We do not use 2014 data to explore longer-run response because the working hours information in 2014 is not directly comparable to that in 2010 and 2012 data. See more details on CFPS data at http://www.isss.edu.cn/cfps/EN/.

larger than 8,000 RMB largely correspond to those facing MTR increases.[71]

Figure A11 shows that for the bottom earners who face no MTR change, working hours on average do not change. For people with wages in 2,000 RMB-8,000 RMB interval, there is evidence that many of them have their working hours increase after the tax reform. But there are also some people having working hours decrease or having no change. For those plausibly facing MTR increases, there is clear evidence that their working hours decrease. Admittedly, self-reported wages and monthly working hours are likely to suffer from measurement errors, figure A11 thus shows suggestive evidence on real responses in working hours. Overall, the working hours change for those facing MTR decreases are mixed and statistically indistinguishable from zero, while that for those facing MTR increases is unambiguous. It is reasonable that higher wage earners are more likely to adjust working hours in response to potential MTR changes, since they are more likely to have the ability or bargaining power to adjust working hours.

Bunching ETI. By contrast to the tax reform ETI, it is more unclear to what extent the bunching ETI captures income manipulation versus real responses. First of all, there is some evidence showing that income manipulation is an important factor behind bunching. For example, Le Maire and Schjerning (2013) derive a dynamic extension to the bunching method, and show that over half of the bunching effect among Danish entrepreneurs is due to intertemporal income-shifting. However, finding clear evidence for bunching ETI reflecting real responses is much harder. Until now, there has been no paper documenting such evidence.[72] Although CFPS survey data provide information on working hours, and we have used it to provide evidence for real responses captured by the tax reform ETI, the self-reported wage is not accurate enough to examine potential real responses around the kinks. Furthermore, the number of observations in the CFPS data are too few to apply the bunching approach. Our data do not allow us to distinguish real responses and income manipulation captured by the bunching ETI. But future research could make progress on this by connecting precisely measured income to working hours.

[71]In particular, with only standard deduction, wages in 12,500-22,000 RMB interval and 38,500-42,000 RMB inverval face an MTR increase (see figure A2). Given potentially high volatility of top earnings, it is not unreasonable to assume most of these top earners face an expectation of increasing MTR.

[72]Chetty et al. (2011) pp. 781-782 try to explore the extent to which income shifting versus labor supply are responsible for the bunching around the top kink. They find no evidence that pension shifting is responsible for a large fraction of the bunching in taxable income at the top kink. But they note that there may be other income shifting responses that they are not able to examine due to the data limitation. Relatedly, Harju and Matikka (2016) estimate real responses and income-shifting responses for the owners of privately held corporations in Finland and find that income-shifting accounts for a majority of overall ETI.

Table A1: Personal income tax on different income components

Item	Description	Tax rate		Standard deduction	
		before reform	after reform	before reform	after reform
1	wages/salaries	9-tier rates	7-tier rates	2000 RMB	3500 RMB
2	income from production or business operation by self-employed industrial and commercial households	5-tier rates	5-tier rates	2000 RMB monthly/ 24000 RMB annually	3500 RMB monthly/ 42000 RMB annually
3	income from the contract and operation of enterprises and institutions and the business income of the lease				
4	remuneration for providing services	20%	20%	4-6: each time when income is received, if income does not exceed 4000, deduct 800, otherwise deduct 20% of the income	
5	author's remuneration	20%	20%		
6	royalties income	20%	20%		
7	interest, dividend, and bonuses	20%	20%	No	No
8	income from leasing property	20%	20%	same as 4-6	
9	property transfer income	20%	20%	original value and reasonable fees	
10	accidental income	20%	20%	No	No
11	other incomes	20%	20%	No	No

Notes: 1. In general, item 1 is taxed monthly, item 2 is taxed yearly, other items are taxed each time the income is received.
2. We call item 2 and 3 self-employment income for simplicity.
3. For item 2, costs, fees, and losses are deductible; for item 3, necessary fees are deductible.
4. Item 4 applies punising rates for excessively high income.
5. Item 5 applies a further 30% exemption of the tax liability.

Table A2: Tax revenues from personal income components

Item		Our sample city (2009.6-2013.12)			China overall (2013)	
		obs	aggregate (RMB)	percent	aggregate (million RMB)	percent
1	wages/salaries	10978979	2222875452	59.05%	409500.31	62.81%
2	income from production or business operation by self-employed industrial and commercial households	0	0	0.00%	57710.38	8.85%
3	income from the contract and operation of enterprises and institutions and the business income of the lease	0	0	0.00%	12181.31	1.87%
4	remuneration for providing services	293572	151349019	4.02%	17412.64	2.67%
5	author's remuneration	162	167059	0.00%	435.84	0.07%
6	royalties income	21	4063418	0.11%	227.05	0.03%
7	interest, dividend, and bonuses	106681	896736376	23.82%	72564.32	11.13%
8	income from leasing property	7	5180	0.00%	2831.06	0.43%
9	property transfer income	2086	453636162	12.05%	67633.07	10.37%
10	accidental income	9617	22296639	0.59%	7746.65	1.19%
11	other incomes	1661	13006288	0.35%	3734.73	0.57%
	Total	11392786	3764135593	100.00%	651977.36	100.00%

Notes: The 2013 national figures come from China Tax Yearbook 2014.

Table A3: Tax schedule of regular monthly wage

pre-reform: standard deduction =2,000 RMB			post-reform: standard deduction =3,500 RMB		
TI	MTR	quick deduction number	TI	MTR	quick deduction number
0-500	5%	0	0-1500	3%	0
500-2000	10%	25	1500-4500	10%	105
2000-5000	15%	125	4500-9000	20%	555
5000-20000	20%	375	9000-35000	25%	1005
20000-40000	25%	1375	35000-55000	30%	2755
40000-60000	30%	3375	55000-80000	35%	5505
60000-80000	35%	6375	>80000	45%	13505
80000-100000	40%	10375			
>100000	45%	15375			

Notes: tax liability is calculated as MTR*TI-quick deduction number.

Table A4: Bounds on underlying elasticity of taxable income with degree of frictions $\delta = 1\%$

taxable income kink	$\Delta\log(1-\tau)$	$\hat{e}$	ε_L	ε_U
pre-reform kink 20000 RMB	-0.06	0.10	0.00	19.41
post-reform kink 9000 RMB	-0.06	0.09	0.00	19.39
post-reform kink 35000 RMB	-0.07	0.41	0.01	17.62
		unified bounds:	$\varepsilon_L^{max} = 0.01$	$\varepsilon_U^{min} = 17.62$

Table A5: Personal characteristics of bunchers v.s. neighboring non-bunchers

	pre-reform 20000		post-reform 9000 RMB		post-reform 35000	
	bunchers	non-bunchers	bunchers	non-bunchers	bunchers	non-bunchers
position	percent	percent	percent	percent	percent	percent
senior chief	5.95	3.89	2.8	1.83	6.03	4.1
senior deputy	1.44	2.88	2.45	2.17	17.68	4.87
middle chief	9.55	4.42	6.11	4.36	6.98	5.76
middle deputy	2.52	2.79	2.2	2.58	6.3	2.55
general staff	78.74	81.75	79.23	80.2	58.2	75.39
other people	1.8	4.27	7.2	8.87	4.81	7.32
occupation	percent	percent	percent	percent	percent	percent
principal of state organs	0.00	0.09	4.31	5.75	1.49	3.16
principal of institutions	0.18	0.35	0.05	0.09	0.14	0.10
principal of enterprises	5.23	3.37	2.85	2.69	7.66	5.55
scientific researcher	0.00	0.20	0.59	0.60	0.07	0.48
engineer or technician	0.18	1.51	5.93	7.06	9.55	4.90
health professional	0.00	0.00	0.19	0.42	0.00	0.14
business personnel	0.18	0.29	0.31	0.29	0.20	0.27
financial personnel	12.97	17.84	15.23	12.80	32.05	11.97
legal personnel	0.00	0.26	0.93	0.06	3.05	0.02
teaching staff	0.36	0.78	1.33	1.67	0.07	0.21
physical education staff	0.00	0.00	0.00	0.00	0.00	0.00
press/publishing and culture industry	0.00	0.00	0.33	0.06	0.07	0.15
clerk	67.21	55.55	49.97	47.92	31.44	55.70
purchasing agent/salesperson	1.44	0.99	0.80	0.72	0.47	0.83
restaurant/tourism/entertainment service person	0.00	0.00	0.06	0.15	0.00	0.09
agricultural industries operating personnel	0.00	0.00	0.01	0.06	0.00	0.01
production/transport equipment operators	4.86	6.33	4.38	4.82	6.84	6.27
other working person	7.39	12.43	12.75	14.84	6.91	10.15
percent of male	64.26	78.2	73.26	75.25	81.55	82.13
mean (s.d.) of age	36.99	41.08	41.28	40.84	42.73	42.98
	(8.88)	(8.85)	(8.92)	(8.96)	(7.94)	(8.15)
observations	555	3442	8080	52520	1496	20503

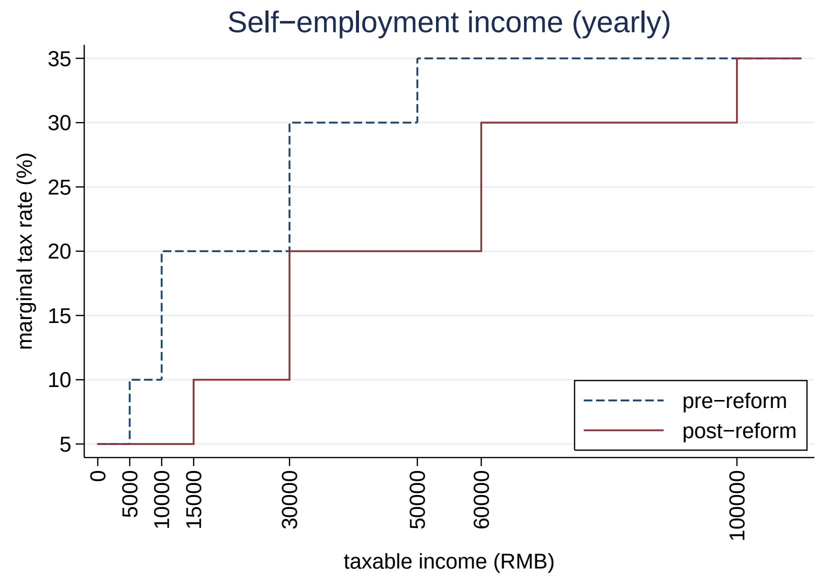
Notes: 1. Some examples for positions are as follows: senior chief-chairman, president, general manage, factory director; senior deputy-general manage, factory director; middle chief-department manager, workshop director; middle deputy-department vice manager, workshop vice director; general staff-people with employment relationship; other people-people without employment relationship or temporarily employed.
2. Bunchers are defined as those falling into bunching region associated with each taxable income kink or standard deduction point, while non-bunchers are those outside the bunching region but close to them as shown in the bunching figures. For an individual falling into bunching region in the figure period, we exclude his/her monthly observations that fall into non-bunching region.

Table A6: Standard tax reform ETI estimates under kink impacts

Dep. Var.: Δlog(taxable income)	[1]	[2]
Δln(1-τ)	2.423***	2.411***
	(0.188)	(0.192)
Δln(1-τ)*around pre-reform kinks		4.452***
		(0.867)
Δln(1-τ)*around post-reform kinks		-1.356***
		(0.398)
Instrument based on	taxable income in t+6	
Observations	1,210,376	1,210,376
First-stage F-stat for Δln(1-τ)	3647	1231
First-stage F-stat for Δln(1-τ)*around pre-reform kinks		210.9
First-stage F-stat for Δln(1-τ)*around pre-reform kinks		479.5

Notes: See notes for table 1.

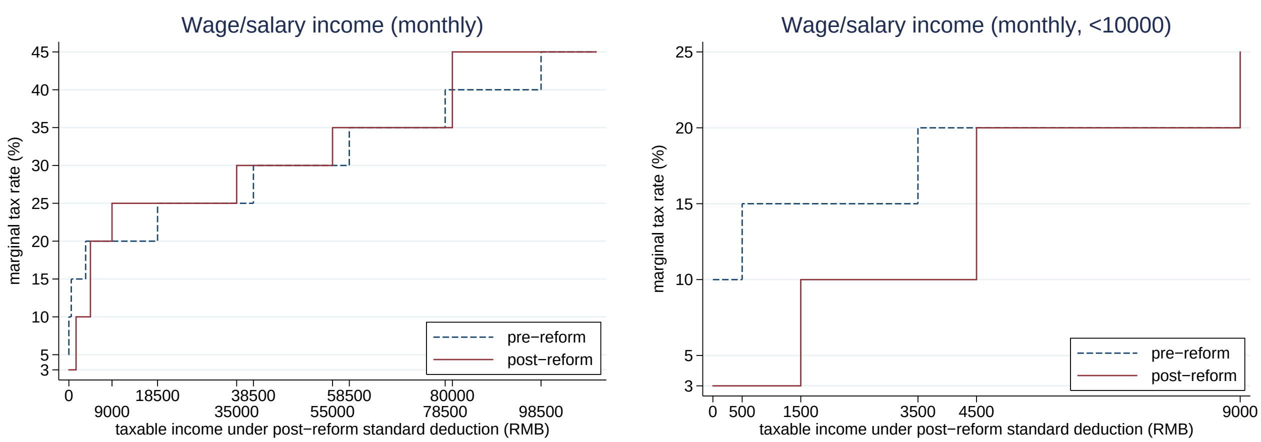
Figure A1: Statutory tax schedule on self-employment income



Notes: The statutory tax schedule applies to very few people with self-employment income in practice. Due to the absence of a reliable accounting book, most self-employed businesses pay a pre-determined fixed amount income tax. More details are discussed in Online Appendix A.

Figure A2: Personal income tax schedule under post-reform tax base

(a) (b)



Notes: Under post-reform tax base, the post-reform standard deduction 3,500 RMB is adopted for both pre-reform and post-reform tax schedules. The new tax rates took effect on September 1st, 2011.

Figure A3: Bunching evidence for full sample

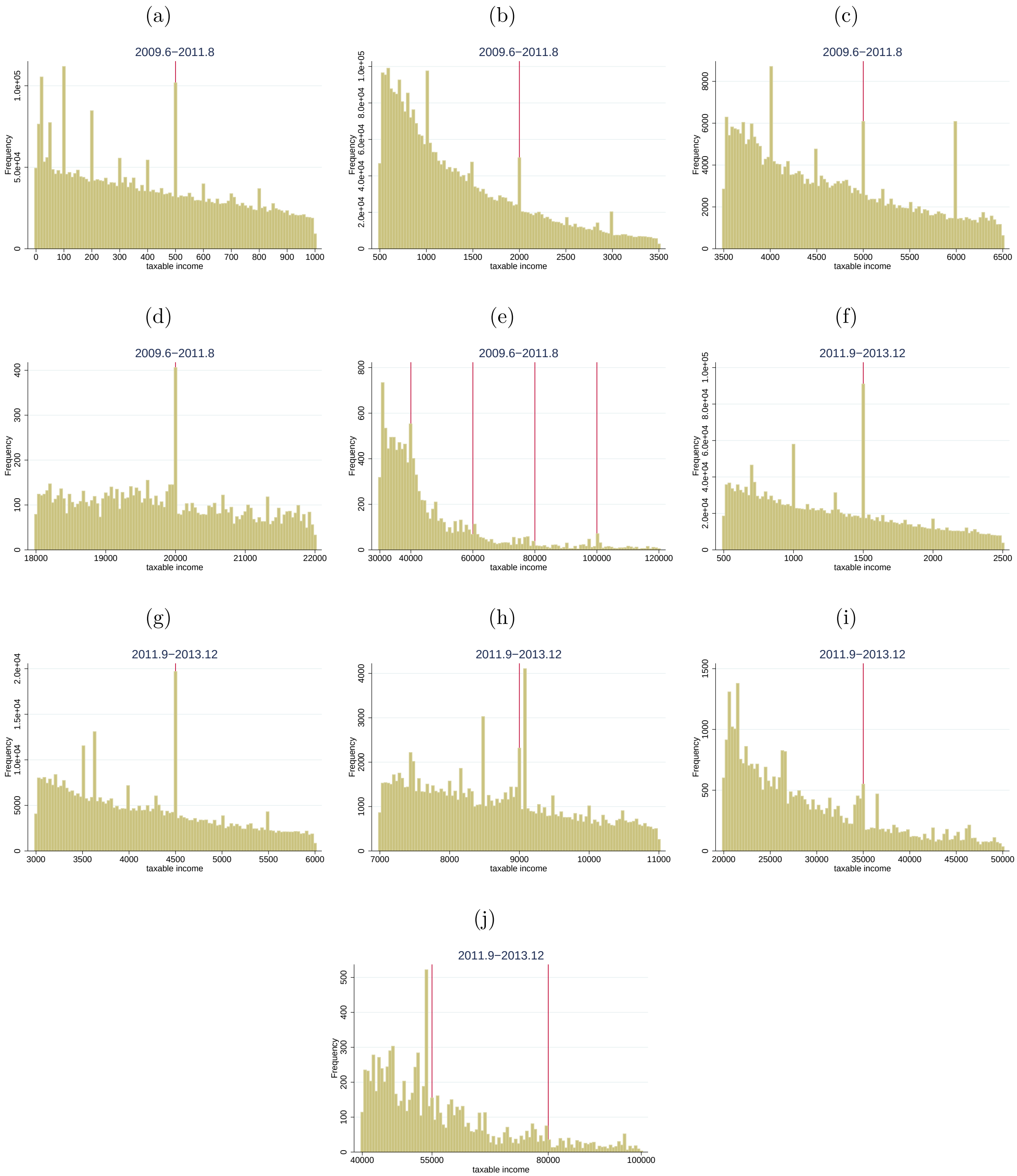


Figure A4: Bunching evidence for decimal sample

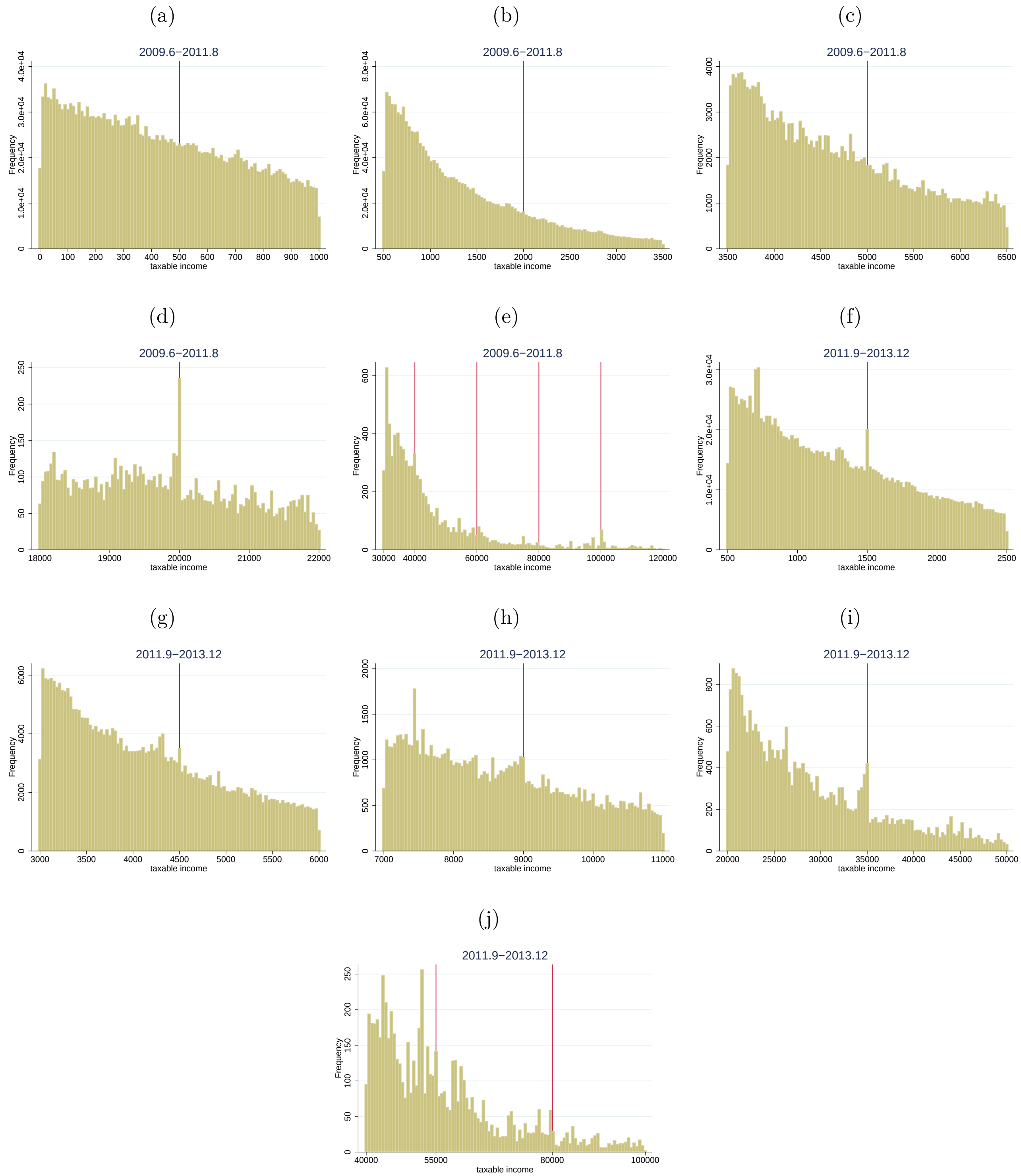
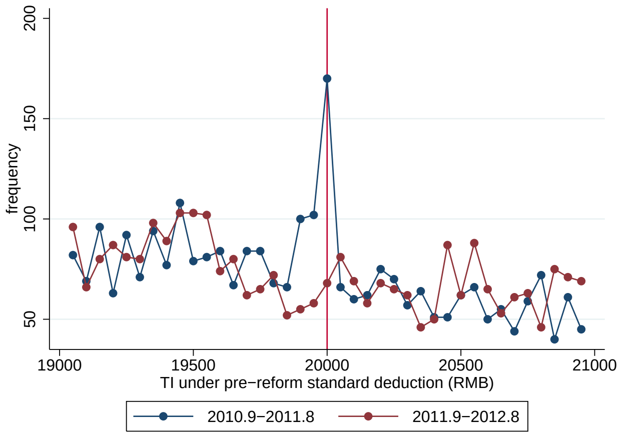
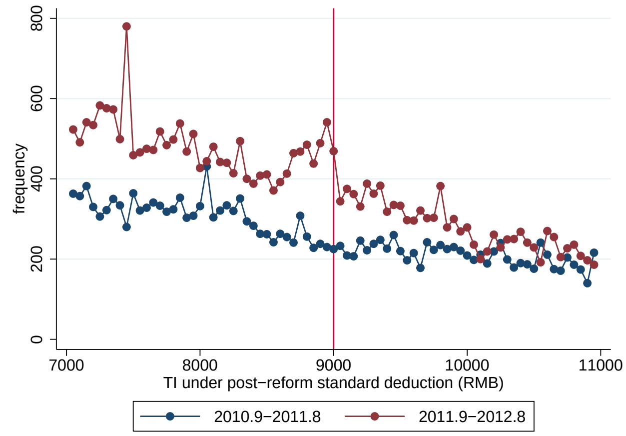


Figure A5: Bunching of taxable income adjusted by standard deduction

(a)



(b)



(c)

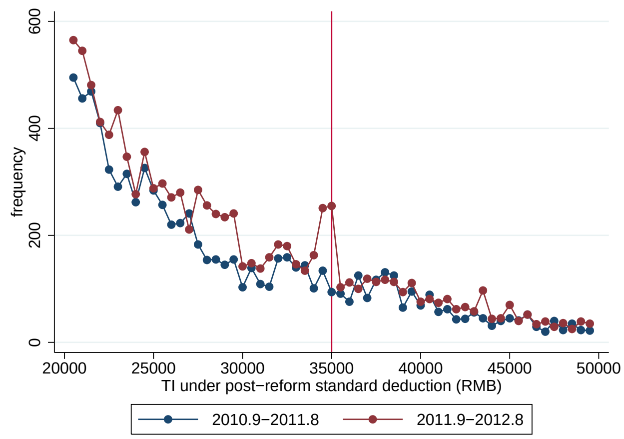


Figure A6: Bunching of unadjusted taxable income

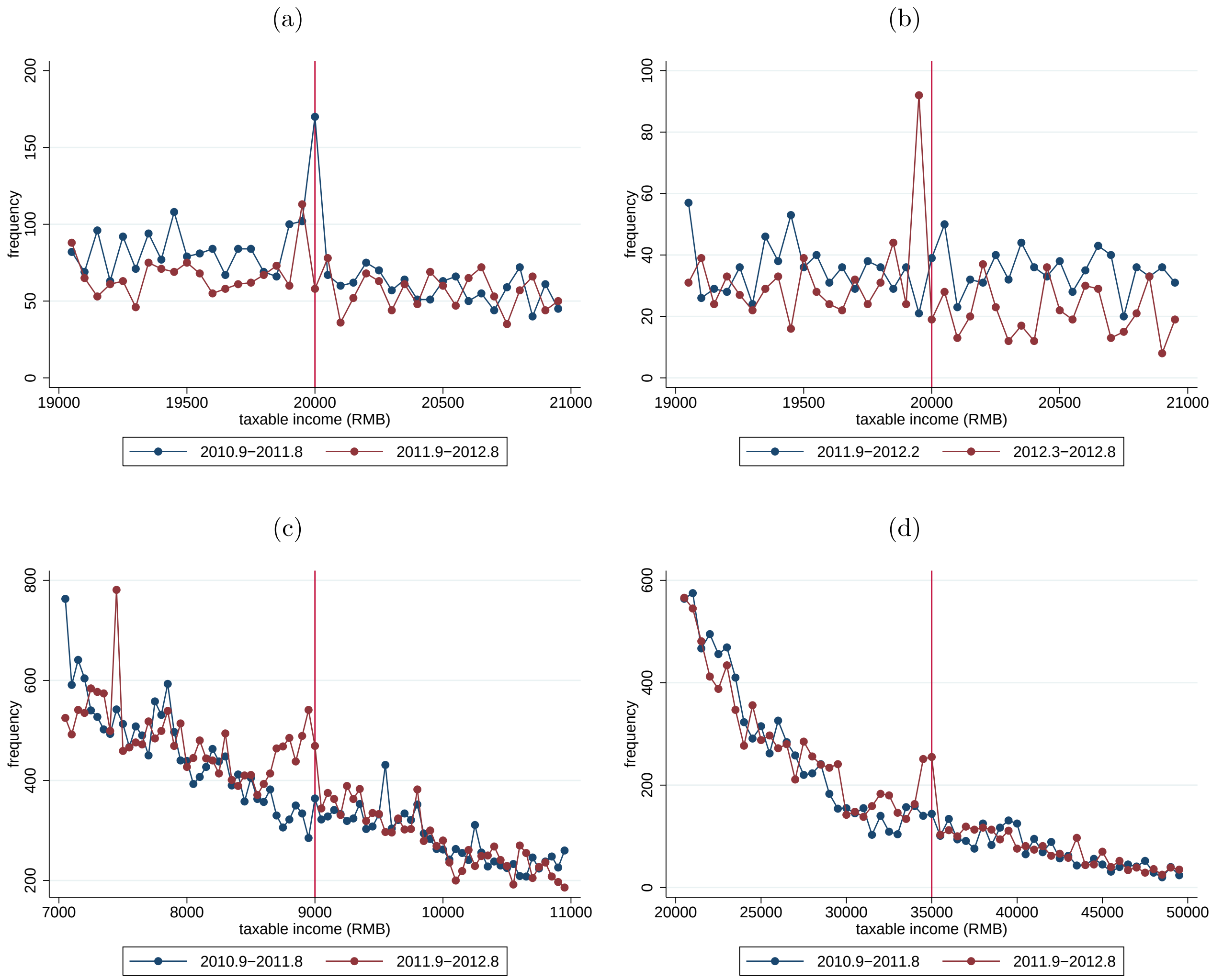


Figure A7: Adjustments of bunching around tax reform

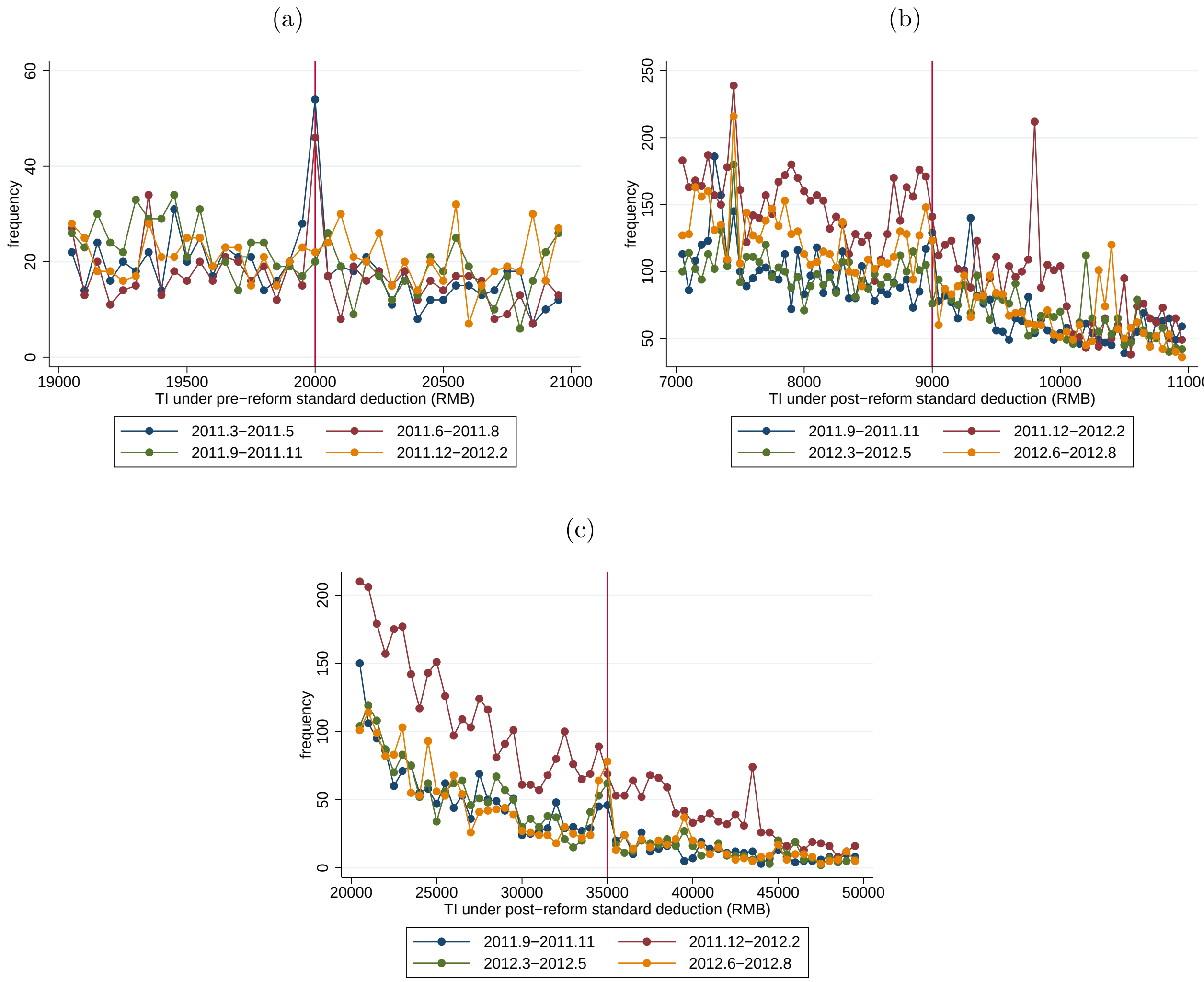
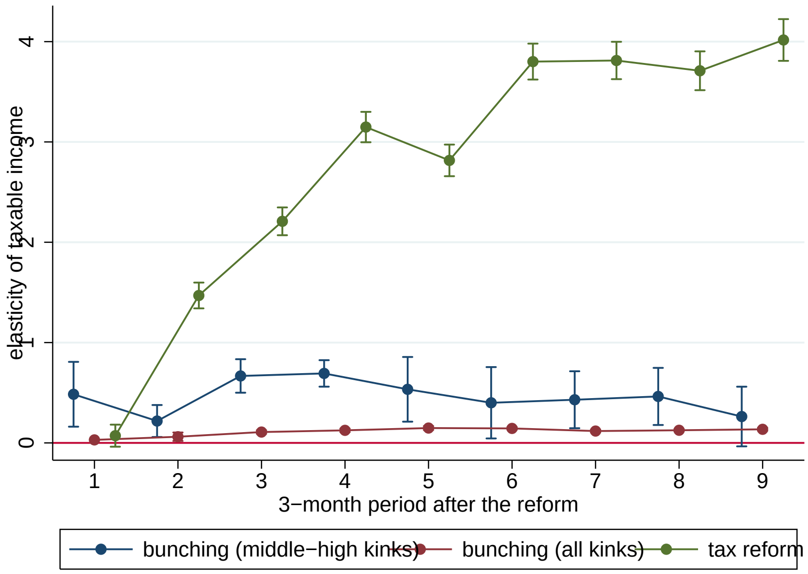


Figure A8: Global ETI estimates over time: tax reform approach vs. bunching approach (full sample)



Notes: A 95% CI of the estimates is shown in the figure. The bunching ETI are estimated using full sample.

Figure A9: Dynamic tax reform ETI estimates: excluding observations around kinks

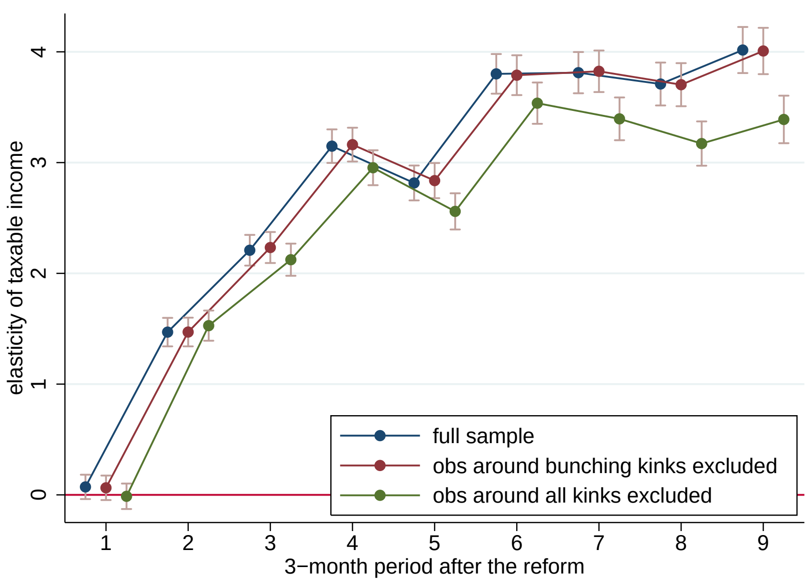
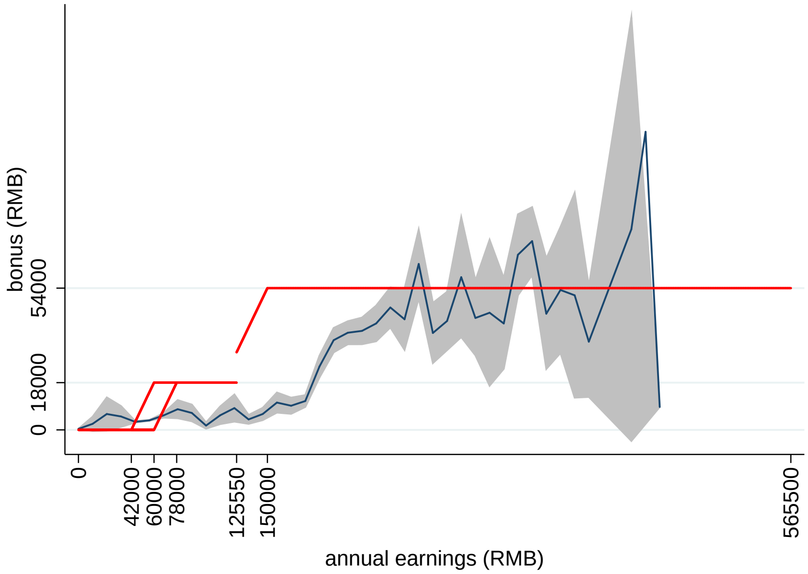


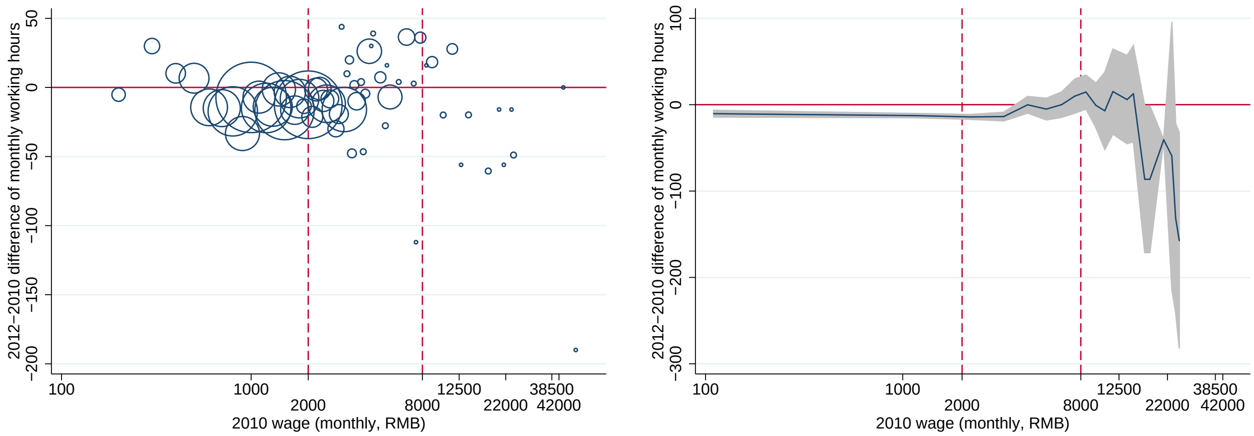
Figure A10: Relationship between bonuses and annual earnings in 2013



Notes: Red line indicates the optimal bonuses policy. Blue line and shadow area indicate fitted polynomial relation between bonuses and annual earnings with a 95% CI.

Figure A11: Responses in monthly working hours to 2011 tax reform

(a) Scatter graph (b) fitted polynomial



Notes: This figure is drawn using 2010 and 2012 CFPS (China Family Panel Studies) data. The sample is restricted to wage earners aged 18-60. In panel a, a bubble indicates mean of change in monthly working hours from 2010 to 2012 in a 100 RMB bin, weighted by the number of individuals in a bin. In panel b, a polynomial is fitted with a 95% CI. The x-axis is shown in log scale. Wages lower than 2,000 RMB face zero tax rate before and after the tax reform. Wages in 2,000 RMB-8,000 RMB interval largely correspond to those facing MTR decreases. Wages larger than 8,000 RMB largely correspond to those facing MTR increases.