CESifo WORKING PAPERS

9773
2022

May 2022

Bunching and Adjustment Costs: Evidence from Cypriot Tax Reforms

Panos Mavrokonstantis, Arthur Seibold

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9773*

Bunching and Adjustment Costs: Evidence from Cypriot Tax Reforms

Abstract

We study adjustment costs in behavioral responses to income taxes, exploiting tax reforms that create and subsequently eliminate income tax kinks in Cyprus. Reduced-form evidence reveals substantial adjustment frictions attenuating bunching and de-bunching responses. Combining the empirical bunching moments with a structural model of frictional earnings supply, adjustment costs are estimated between EUR 93 and EUR 238 for wage earners. Moreover, we uncover important asymmetries in adjustment frictions, where bunching at a kink is costlier than de-bunching away from the kink. Finally, we find that self-employed individuals face considerably lower adjustment costs than wage earners.

JEL-Codes: H240, J220.

Keywords: income taxation, taxable income responses, bunching, adjustment frictions.

Panos Mavrokonstantis
Department of Economics
London School of Economics / United Kingdom
p.mavrokonstantis@lse.ac.uk

Arthur Seibold
Department of Economics
University of Mannheim / Germany
seibold@uni-mannheim.de

April 2022
We are grateful to Michel Azulai, Marcos Charalambides, Sarah Clifford, Frank Cowell, Giulia Giupponi, Dana Kassem, Henrik Kleven, Camille Landais, Alexandra Patsalides-Charalambides, Viktorie Sevcenko, Johannes Spinnewijn, Eddy Tam, three anonymous referees, and numerous seminar and conference participants for helpful comments and discussions. We especially thank Christian Hilscher who provided outstanding research assistance for this project. We would also like to thank the Tax Department of the Republic of Cyprus for providing access to their administrative data, in particular Elina Symeonidou who prepared and explained the data. The views expressed here are those of the authors and do not necessarily reflect the views of the Tax Department of the Republic of Cyprus.

1 Introduction

Canonical models of labor supply assume that workers can adjust their earnings in a frictionless manner. This view is challenged by mounting evidence that workers face significant optimization frictions in responding to tax policy changes (Chetty et al., 2011; Kleven and Waseem, 2013; Gelber et al., 2020). Frictions may reflect a variety of factors attenuating short-term adjustments, including the cost of renegotiating contracts or of searching for new jobs, inattention to reforms, or a lack of knowledge of tax incentives. Incorporating adjustment costs can explain several empirical puzzles in public finance, such as the large difference between estimates of micro and macro labor supply elasticities (Chetty, 2012; Chetty et al., 2013) and the lack of observed responses to jumps in marginal tax rates at some income tax kinks, in particular by wage earners (Saez, 2010; Chetty et al., 2011). Moreover, the presence of frictions has implications for optimal tax policy and welfare evaluation (Werquin, 2016; Farhi and Gabaix, 2020).

In this paper, we study adjustment frictions in behavioral responses to personal income taxes, using administrative tax data from Cyprus and quasi-experimental variation in the location of kinks in the tax schedule. We combine reduced-form bunching techniques with structural methods in order to estimate earnings responses and adjustment costs, building on the methodology proposed by Gelber et al. (2020). Our paper makes two main contributions. First, we document sizeable bunching and provide evidence of adjustment frictions among wage earners, a group for whom such clear responses are rarely observed in the literature. Second, we find that adjustment costs attenuating earnings responses to income taxes are substantial, and provide novel evidence of asymmetry in such frictions. In particular, our results suggest that bunching precisely at a kink is costlier to workers than de-bunching away from the kink.

Our empirical setting is given by the income tax system in the Republic of Cyprus. This setting has some important advantages for our purposes. First, a number of tax reforms in the 2000s provide exogenous variation in the location of tax kinks, i.e. discontinuities in marginal tax rates. In particular, the setting enables us to track the full trajectory of earnings responses from the time a kink is introduced, while it remains in place, and after it is removed. Second, the Cypriot income tax system closely resembles typical tax systems in other E.U. countries and the U.S., making our results comparable to the broader literature on taxable income responses. Third, high-quality administrative data is available for our analysis. We use a novel dataset of administrative tax records spanning the years 2003 to 2010, providing more than one million observations.

We begin by presenting several reduced-form findings. We obtain compelling graphical evidence of earnings responses among wage earners by estimating bunching around various kinks in the income tax schedule. In particular, we find strong bunching at the first tax kink, whereas patterns around kinks at higher income levels are more mixed. Furthermore, we provide two pieces of reduced-form evidence of adjustment frictions by examining the taxable income distribution during tax reforms. First, when a new kink is introduced, bunching builds up gradually, suggesting that some workers are constrained by frictions that prevent them from responding immediately. Second, we show that some workers continue to bunch at former kinks which are removed by tax reforms. These findings stand in contrast to frictionless models, which would predict that bunching appears immediately at a new kink and de-bunching is immediate and complete upon the removal of the kink. Moreover, we document the dynamics of bunching and de-bunching responses. Bunching at a new kink begins in the first year but keeps growing for three years after its introduction, while it takes two years for residual bunching to be completely eliminated from a former kink.

Next, we specify a model that allows us to translate the reduced-form evidence into estimates of the elasticity of taxable income and a fixed cost of adjusting earnings. The model yields indifference conditions of marginal individuals who are indifferent between remaining at their initial earnings and incurring adjustment costs to bunch at the kink, or conversely between remaining at the kink and de-bunching away from the kink. We first consider the standard model following Gelber et al. (2020), assuming that bunching originates from above the kink, as is usually the case in the bunching literature. We also present an adapted model, where bunching originates from below the kink. The adapted model explicitly takes into account that kinks are created by shifting bracket thresholds upwards in our empirical setting, a situation often encountered in the context of income tax reforms. Exploiting the panel dimension of our data, we provide additional suggestive evidence supporting the notion that bunching at new kinks can originate from below. However, the adapted model is less well in line with observed de-bunching responses, which tend to take individuals further away from the kink than predicted if they originally bunched from below. Finally, we extend the models to dynamic scenarios where individuals face adjustment frictions with some probability in each period after a tax reform.

We present structural parameter estimates based on both model variants. Our preferred specifications reveal sizeable adjustment costs for wage earners between CYP 54 (EUR 93) and CYP 140 (EUR 238). On average, these costs represent around 0.5% of 1.5% of taxable income among affected individuals. Moreover, our results suggest important asymmetries in adjustment frictions, where bunching exactly at a kink entails greater costs than de-bunching away from a kink. De-bunching costs are estimated at only CYP 22 (EUR 37) to CYP 44 (EUR 74) in the standard model, and close to zero in the adapted model. Estimates of the elasticity of taxable income vary between 0.04 and 0.29, depending on the set of kinks used in the estimation. The dynamic model estimation implies similar magnitudes of adjustment costs and elasticities, and provides additional information on the speed of adjustment. Consistent with the reduced-form evidence, the estimated probability to face adjustment frictions declines over time. In the first year after a tax reform, up to 67% of workers are constrained by frictions, but the probability approaches close to zero by the third year.

While our main analysis concentrates on wage earners, we also consider bunching and adjustment frictions among the self-employed. In contrast to wage earners, who bunch mainly at the first tax kink, the clearest bunching responses among self-employed individuals can be found around the top tax kink. Again, we find compelling reduced-form evidence of adjustment frictions, with bunching building up gradually over three years, and some residual bunching after the kink is removed. However, bunching is sharper and of larger magnitude than among wage earners. Repeating the structural estimation for the self-employed suggests that these differences can be explained by smaller adjustment costs of only CYP 7 (EUR 12) to CYP 22 (EUR 38). Moreover, the self-employed are characterized by large structural elasticities between 0.30 and 0.55.

This paper contributes to a nascent literature combining reduced-form bunching techniques (see Kleven (2016) for a review) with structural estimation of adjustment costs. In their seminal work, Gelber et al. (2020) estimate adjustment costs in the context of the U.S. Social Security earnings test. A closely related paper by Zaresani (2020) estimates adjustment costs among disability insurance recipients in Canada. We make two main contributions to this literature. First, exploiting a setting with multiple tax reforms, we document the full trajectory of bunching dynamics from before a kink is created until after it is removed, which enables us to provide novel evidence of asymmetries in adjustment frictions. In particular, our results suggest that when adjustment requires locating at a specific earnings level, such as bunching at a kink, this entails greater frictions than moving away from this earnings level. Second, we estimate adjustment costs in a standard income tax setting, complementing the evidence from social security settings considered in previous work. Although taxable income responses represent the most prevalent context of bunching estimation in the literature, there are very few existing estimates of adjustment costs from such settings. A notable exception is given by He et al. (2021), who study taxable income responses in China and find very small adjustment costs of RMB 1.22 (USD 0.19).[1] In contrast, we provide evidence of sizeable adjustment costs in taxable income responses, and our preferred estimates are more in line with magnitudes from previous work, including the estimate of USD 280 from Gelber et al. (2020).

More generally, our paper is related to the extensive literature estimating the elasticity of taxable income (see e.g. Saez et al. (2012) for a review). One set of studies use tax reforms to quantify taxable income responses (e.g. Feldstein 1995; Goolsbee 2000; Gruber and Saez 2002; Blomquist and Selin 2010; Gelber 2012; Kleven and Schultz 2014), while others use bunching methods (e.g. Saez 2010; Chetty et al. 2011; Kleven and Waseem 2013; Bastani and Selin 2014). In contrast to much of the literature, we find clear evidence of bunching at tax kinks among wage earners. To our knowledge, we provide the first comparison of adjustment costs among wage earners and self-employed individuals. Our results confirm conjectures that the typically large observed bunching responses of the self-employed can be explained by less severe adjustment frictions. Finally, while much of the literature focuses on the United States or Scandinavian countries, the Cypriot setting provides insights into behavioral responses to income taxes in

---

[1] Although not the main focus of their paper, Gudgeon and Trenkle (2021) also provide estimates of adjustment costs among marginally employed women in Germany.

a new context: that of a small open economy, with an intermediate level of GDP per capita compared to other developed economies.[2]

The remainder of the paper proceeds as follows. Section 2 describes the institutional context and data, Section 3 presents reduced-form evidence of bunching responses and adjustment costs, Section 4 sets out the theoretical framework, Section 5 discusses the structural estimation results, Section 6 shows additional results for the self-employed, and Section 7 finally concludes.

2 Institutional Context and Data

2.1 The Cypriot Income Tax System

Cyprus has a simple personal income tax system, sharing many of its features with typical tax systems in other European countries and the U.S. Marginal tax rates are progressive and there is a single measure of taxable income. Taxable income is defined as the sum of individual income from employment, business income, pension income, foreign income and capital income, net of deductions. The main deductions include social insurance contributions, pension fund contributions, life and medical insurance premia, trade union subscriptions, charitable donations, and various types of capital income deductions. Both wage earners and the self-employed are subject to the same tax schedule, which is not indexed to inflation. The local currency until 2007 was the Cypriot Pound (CYP), after which it was replaced by the Euro at an exchange rate of around EUR 1 $\approx$ CYP 0.585.

All wage earners with gross income above the personal allowance must file a tax return.[3] In practice, many individuals earning below the exemption threshold also file because it is a requirement for access to government welfare programs, or employers file on their behalf. Individuals earning any income from self-employment must always file, irrespective of the amount earned. The enforcement system involves third-party reporting for employees, and tax withholding for both employees and self-employed. Tax returns are filed individually.

Figure 1 shows marginal income tax rates throughout our period of analysis. Due to its progressive rate structure, the Cypriot income tax system creates convex kinks in workers' budget sets. The largest jump in the marginal tax rate occurs at the first kink, where it increases from zero to 20%. Marginal tax rates then rise to 25% at the second kink and to 30% at the third kink. Moreover, a series of tax reforms in 2004, 2007 and 2008 provide rich variation in the location of kinks. Until 2003, the tax bracket cutoffs are located at CYP 9,000, CYP 12,000 and

---

[2] We contribute to the limited existing work on the Cypriot labor market. Previous studies focus on estimating labor supply functions (Pashardes and Polycarpou, 2010) and examining the determinants of labor supply among welfare recipients (Pashardes and Polycarpou, 2012).

[3] The personal allowance is nominally equivalent to the first tax bracket cutoff below which a zero marginal tax rate applies. It is important to note, however, that the tax filing threshold is based on gross income, while the tax schedule itself is based on taxable income. Virtually all tax filers have some deductions, such that their gross income is larger than taxable income. In Section 3.2, we discuss this issue further and provide evidence that the filing threshold does not drive observed bunching at the first tax kink.

CYP 15,000.[4] In 2004, new kinks are created at CYP 10,000, CYP 15,000 and CYP 20,000 by shifting the location of the respective cutoffs. These kinks stay in place for three years between 2004 and 2006, and are eliminated in 2007, when tax bracket cutoffs are moved again. The kinks created in 2007 at CYP 10,750, CYP 15,750 and CYP 20,600 stay in place for only one year, after which the cutoffs are shifted to CYP 11,413, CYP 16,388 and CYP 21,246 in the course of the introduction of the Euro. The main goal behind these repeated tax reforms was to adapt the tax schedule to relatively fast earnings growth Cyprus experienced in the early 2000s, avoiding bracket creep.[5]

These reforms provide variation ideally suited for our purposes. In particular, the reforms of 2004 and 2007 enable us to trace the full trajectory of responses to a kink, including the building-up of bunching at a newly created kink, as well gradual de-bunching at the same kink once it is removed. The other tax reforms provide additional bunching observations across different earnings levels, varying in the time since the creation or removal of the respective kink. Finally, the yearly bunching estimates are easily comparable, as there are no significant changes to the tax base during our period of analysis from 2003 to 2010.

2.2 Data

We obtained first-time access to an administrative dataset covering the universe of tax filers in the Republic of Cyprus. The dataset contains information from the main tax return form (the IR1A), as well as basic demographic and firm characteristics. We use data covering the years 2003 to 2010, and restrict our estimation sample to working individuals aged 25 to 54 in order to avoid confounding effects from part-time work or early retirement. We focus on two types of workers: wage earners, who only have earnings from employment, and the self-employed, who only have earnings from self-employment. We do not consider a very small proportion of 1% of the sample that has income from both salaried employment and self-employment. These restrictions result in a sample of 1,044,808 observations.

Summary statistics of the data are shown in Table 1. On average, tax filers earn CYP 14,332 of annual gross income and receive CYP 1,631 of deductions, leading to taxable income of CYP 12,703 and a tax bill of CYP 891 per year. Females comprise about 40% of the sample, the average age is 40.7, and about 91% are wage earners. Finally, the largest sector of activity is the public sector, followed by services, trade, finance, manufacturing and construction.

---

[4] The kinks at CYP 9,000 and CYP 12,000 were in place since 2000, while the kink at CYP 15,000 was created in 2003. We do not use the years before 2003 because different rules for deductions applied during this time.

[5] The tax reforms were not officially announced to be of a temporary nature, but taxpayers could be aware that bracket cutoffs may be shifted repeatedly based on prior experience. Yet, we do not find evidence of anticipatory behavior in responses to these kinks in Section 3.

3 Reduced-Form Evidence of Adjustment Frictions

3.1 Bunching Estimation

In order to obtain reduced-form evidence of responses to kinks in the tax schedule and adjustment frictions, we use the bunching method introduced by Saez (2010) and Chetty et al. (2011). To measure the bunching mass B at a taxable income threshold z^*, we have to estimate what the counterfactual distribution $h_0(z^*)$ would have been in the absence of the kink, and compare this to the observed distribution in the presence of the kink. B is the difference between the observed spike in the distribution and the counterfactual density of taxable income. To make responses comparable across kinks, the bunching mass can be normalized relative to the counterfactual, resulting in the excess mass b. Finally, the bunching response can be linked to the kink size $\Delta\tau$, computing an observed elasticity of taxable income w.r.t. to the net-of-tax-rate $1-\tau$[6]

$$\hat{\varepsilon} = \frac{\Delta z^*/z^*}{\Delta\tau/(1-\tau)}$$

where Δz^* is the earnings response of the marginal bunching individual, which can be inferred from estimated bunching via $\Delta z^* \approx B/h_0(z^*)$. Note that the observed elasticity $\hat{\varepsilon}$ calculated according to the Saez (2010) method can under-estimate the true structural elasticity in the presence of adjustment frictions (see Section 4).

In practice, we group individuals into earnings bins of size 50, and we fit a flexible 9th-order polynomial to the empirical distribution of taxable income to obtain counterfactual distributions. Because many kinks are located at round numbers, and those may act as reference points, we control for round-number fixed effects. Furthermore, we include fixed effects for bins in the set of tax bracket cutoffs in the estimation bandwidth, thus netting out any influence of "nearby" kinks from the counterfactual. The bunching region around each kink is visually determined, allowing for some potential diffuse bunching. Standard errors are bootstrapped. We also show results under alternative choices of key estimation parameters, and Appendix C.1 presents details of the bunching estimation.

3.2 Reduced-Form Bunching Evidence

In this section, we present reduced-form evidence of behavioral responses to tax kinks and adjustment frictions. Our main analysis focuses on wage earners, and Section 6 shows additional results for the self-employed. We begin by examining the CYP 10,000 kink, where the full trajectory of bunching responses can be observed. The kink is introduced in 2004, remains in place for three years until 2006 and is removed in 2007. Thus, a model without frictions would predict no bunching in 2003, stable bunching from 2004 to 2006, and complete, immediate

---

[6] As long as $\Delta\tau$ is small, this represents a compensated elasticity. With large kinks, the estimate is a weighted average of compensated and uncompensated elasticities (Kleven, 2016).

de-bunching in 2007.

Figure 2 shows the empirical distribution of taxable income and the estimated counterfactual around the CYP 10,000 kink for the years 2003 to 2008. Each panel reports the estimate of the excess mass b, and provided that the kink is in place in that year, the estimate of the observed elasticity $\hat{\varepsilon}$. Panel (a) shows that the distribution is smooth around CYP 10,000 in 2003, suggesting no anticipatory behavior in the year before the kink is introduced. In 2004, the first year when the tax schedule features a kink at CYP 10,000, a clear spike at the location of the kink emerges (Panel b). The estimated excess mass is 1.04, implying that there are 104% more workers locating at the kink than what would be expected under the smooth counterfactual distribution. Bunching grows in the following years, with the excess mass increasing to 1.16 in 2005 (Panel c) and 1.78 in 2006 (Panel d). As the size of the kink remains constant across years, this implies an increase of the observed elasticity from 0.026 to 0.044. This gradual building-up of bunching over three years provides a first piece of evidence of adjustment frictions, as some workers do not seem to be able to adjust their earnings towards the kink immediately.

A second, striking piece of evidence of adjustment frictions is provided by the 2007 distribution shown in Panel (e) of the figure. There is significant bunching at CYP 10,000 ($b = 0.49$), although there is no tax kink at this earnings level any more. The reduction in the excess mass compared to 2006 implies that many workers adjust, but the presence of significant residual bunching suggests that some workers face adjustment frictions preventing them from immediate de-bunching. In 2008, the second year in which the kink is removed, there is no more significant residual bunching at the kink, although the point estimate of the excess mass remains positive.

Figure 3 shows bunching responses to other kinks during the sample period. To begin with, Panels (a) to (f) show taxable income distributions around the CYP 9,000 and CYP 12,000 kinks, which are in place in 2003 and removed in 2004. At both kinks, there is significant excess mass of 0.31 and 1.04 in 2003, respectively, implying observed elasticities between 0.01 and 0.07. The CYP 12,000 kink in particular provides additional evidence of adjustment frictions, as substantial residual bunching of 1.07 can be observed in 2004, the year after the kink is removed. Even in 2005, the second year after its removal, the excess mass at the former kink is still 0.65. On the other hand, we do not detect significant residual bunching at the CYP 9,000 kink after it is removed. Next, the CYP 10,750 kink is in place only in 2007. Panel (h) shows sizeable excess mass of 1.50 at this kink, corresponding to an observed elasticity of 0.03. Residual bunching in 2008 is positive but not significant in Panel (i). Finally, Panels (j) to (l) show bunching at the CYP 11,413 kink, which is introduced in 2008 and remains in place until the end of the sample period. We estimate large excess mass between 1.94 and 2.31 and observed elasticities of 0.04 to 0.05 at this kink. Overall, these sharp observed bunching responses provide compelling evidence of behavioral responses to tax kinks among wage earners, a group for whom such clear bunching is rarely found in the literature (Saez, 2010; Bastani and Selin, 2014). Moreover, the results suggest that adjustment frictions are most clearly visible at kinks which stay in place sufficiently long and where initial bunching is large.

In addition, Appendix Figure A1 shows the distribution of taxable income around all remaining tax kinks, including the kink at CYP 15,000 in the years 2003 to 2006, the CYP 20,000 kink from 2004 to 2006, the kinks at CYP 15,750 and CYP 20,600 in 2007, and the kinks at CYP 16,388 and CYP 21,246 from 2008 to 2010. There are generally no significant observed bunching responses around these kinks at higher earnings levels, except at the CYP 15,000 in some of the years while it is in place. The finding of bunching at some kinks but no bunching at others mirrors results from the literature, such as Saez (2010) who finds bunching only at the first kink of the U.S. income tax schedule.[7]

Several factors could potentially explain the lack of bunching at higher earnings levels in our setting. First, the elasticity of taxable income among individuals in the neighborhood of these kinks could be very low. However, this explanation would be contrary to existing evidence that high income earners tend to respond more elastically to tax changes (Saez et al., 2012). Second, higher-income kinks feature smaller changes in marginal tax rates than the first kink, which could make incurring adjustment costs to bunch less worthwhile. Third, adjustment frictions could be larger among higher earners. One possible reason for larger frictions may be that many higher-paying jobs in Cyprus are in unionized sectors with more rigid working arrangements, including the large public sector. Lower earners, on the other hand, often work in sectors such as services (including tourism), construction and trade, where seasonal work and other more flexible arrangements are more common.[8] Empirically, we cannot separately identify a low elasticity from adjustment costs in the absence of any bunching response. Thus, our subsequent analysis focuses on the kinks shown in Figures 2 and 3, where significant bunching responses are observed while they are in place.

Robustness of Bunching Estimates. Appendix Table A1 investigates the robustness of our bunching estimates to alternative estimation specifications, focusing on the case of the CYP 10,000 kink.[9] The excess mass and observed elasticity estimates in each year remain similar when considering alternative choices of estimation parameters, including the bandwidth around the kink, the size of taxable income bins, the extent of the excluded region around the kink, and the polynomial order of the counterfactual estimation. Moreover, we consider various alternative strategies to estimate the counterfactual density. On the one hand, we compute the counterfactual as a polynomial like in the baseline estimation, but with an additional shift of the density above or below the kink. These corrections take into account that bunching may originate from the right or from the left of the kink in our setting (see Section 4). Reassuringly, shifting

---

[7] Similarly, Chetty et al. (2011) find that wage earners only bunch at one kink in Denmark, although it is the top tax kink in their case.

[8] Appendix Table A2 shows summary statistics for wage earners around the first kink vs. higher earners, illustrating these sectoral differences. Unfortunately, we cannot investigate the role of working arrangements more directly because our data is at the annual level and does not include information on hours of work.

[9] Appendix B provides similar robustness results for the full set of kinks used in the analysis.

the density on either side does not substantially change our estimates. On the other hand, we implement a variant of the counterfactual based on the empirical distribution in the year before the introduction of the kink, rather than fitting a polynomial. This strategy exploits the panel dimension of our data and addresses the Blomquist et al. (2021) critique of the identification of bunching responses in cross-sectional data.[10] If anything, this procedure somewhat increases estimated bunching, while dynamic patterns remain similar.

Role of the Tax Filing Cutoff. As mentioned in Section 2.1, only individuals with gross income above the personal allowance are required to file a tax return. This may raise concerns about responses to the first tax kink being driven by the filing requirement. However, it is important to note that the filing threshold is based on gross income, while the tax schedule is based on taxable income. Virtually all tax filers have some deductions, such that their gross income exceeds taxable income. On average, individuals' gross income is 12.8% higher than their taxable income (see Table 1). Appendix Figure A2 confirms that most bunching at the first tax kink is not driven by the filing cutoff, focusing on the CYP 10,000 kink. In Panel (a), gross incomes indeed spike at the filing threshold, but at the same time many taxpayers file a tax return even if their gross income is below the threshold. Crucially, Panel (b) shows that 92% of bunchers at the first tax kink have gross income above CYP 10,000, such that bunching at the kink does not coincide with being located at the filing threshold. There is also no sizeable discontinuity in the probability of being located above the filing threshold around the CYP 10,000 kink in Panel (c). Finally, Appendix Figure A3 shows bunching patterns around the CYP 10,000 kink, excluding those individuals who are located at the filing cutoff. Excess mass and observed elasticity estimates remain similar to those in Figure 2, further corroborating that bunching at the first tax kink is not substantially confounded by the filing threshold.

3.3 Mechanisms Behind Bunching Responses

Next, we examine the mechanisms behind taxable income adjustments in response to the creation or the removal of a tax kink. Figure 4 presents evidence on the dynamics of individual bunching and de-bunching at the CYP 10,000 kink. Panel (a) considers individuals who bunch at the kink for the first time in any of the years while it is in place, and shows their taxable income distribution in the year before they bunch. Bunching originates from income levels both below and above the kink, but the majority of 79% of bunchers were located below the kink in the prior year. This may seem surprising from the point of view of standard bunching models where bunching is typically assumed to come from above the kink. However, new kinks are created by shifting existing bracket cutoffs upwards in our setting, which implies that individuals who had

---

[10] To account for nominal taxable income growth throughout our sample period, we adjust the 2003 distribution for inflation and normalize for the relative sample size within the bandwidth in each year. See Appendix Figure A4 for bunching plots under this alternative counterfactual.

already responded to the old kink can bunch from below at the new kink. We return to this conceptual discussion in detail in Section 4.

Panel (b) of Figure 4 shows previous-year taxable income of residual bunchers in 2007, the year the kink is removed. The distribution features a very large spike at the kink, suggesting that most residual bunching is driven by inertia of individuals who bunched at the kink while it was in place. This suggests that adjustment frictions are indeed responsible for bunching observed after the kink is removed. Furthermore, Panel (c) provides evidence on the direction of de-bunching responses, focusing on individuals who bunched in the last year the kink is in place. Again, there is strong residual bunching in the following year, but the figure also suggests that most de-bunching adjustments occur by individuals increasing their taxable income above the kink. Moreover, the shape of the distribution among former bunchers to the right of the kink yields additional insights into de-bunching responses. The density is relatively low just above the kink and increases steeply thereafter. This suggests that most former bunchers have counterfactual earnings substantially above the kink, to which they revert when the kink is removed. The de-bunching patterns are in line with a sizeable fixed cost of adjusting earnings, which only makes it worthwhile to adjust if the change in earnings after de-bunching is sufficiently large.

Figure 5 provides additional evidence on the correlation of bunching with growth in taxable income, gross earnings and deductions. Panel (a) splits the sample into individuals experiencing growth vs. reductions in their taxable income relative to the prior year. Bunching at the CYP 10,000 kink is estimated separately for each subgroup, pooling the years while the kink is in place. The absolute magnitude of spikes at the kink among the two groups indicates that most bunching is driven by individuals with growing income. Nevertheless, there is also clear bunching relative to the local density among workers who reduce their taxable income, and the excess mass among this subgroup is 0.67, compared to 1.04 among those with income growth.

In addition, Panels (b) and (c) of Figure 5 shed light on the anatomy of bunching responses, examining whether bunching responses are associated with particular changes in gross earnings or deductions. In Panel (b), most bunching at the kink is associated with growth in gross earnings. The respective excess mass estimates of 1.09 and 0.95 among workers with gross earnings growth vs. reductions are relatively similar to those by taxable income growth. This is consistent with wage earners mostly bunching via adjusting gross earnings. Panel (c) shows a smaller difference in bunching between workers with growing vs. reduced deductions, and the excess mass estimates of 0.76 and 0.71 among the two groups are more similar. Thus, the role of adjusting deductions in bunching responses is less clear. Since deductions tend to be easier to manipulate than third-party reported salary earnings, these results have potential implications for whether bunching captures real or reporting responses. In particular, the association of bunching with gross earnings adjustments is consistent with real responses to changes in marginal tax rates.

4 Conceptual Framework

4.1 Basic Bunching Framework

We first fix ideas by presenting a simple model of earnings supply without frictions following Saez (2010) and Kleven (2016). Workers have preferences over consumption c and taxable income z, and differ in their ability n. Utility $u(c, z; n)$ is increasing and concave in consumption ($u_c > 0$, $u_{cc} < 0$), and decreasing and convex in earnings ($u_z < 0$, $u_{zz} < 0$), capturing disutility of work. Moreover, $u_{zn} < 0$, i.e. the marginal disutility of generating earnings falls with ability. Individual ability is the only source of heterogeneity and is distributed according to a smooth density function $f(n)$. Individuals pay taxes on income z captured by $T(z)$, which may be a nonlinear function. Denoting virtual income by y,[11] the worker's optimization problem can be stated as:

$$\max_{c,z} \quad u(c, z; n) \quad s.t. \quad c = (1-\tau)z + y$$

The first-order condition $(1-\tau)u_c + u_z = 0$ implicitly defines the optimal earnings supply function $z = z((1-\tau), y; n))$. Under a smooth ability distribution, and a linear tax schedule with constant marginal tax rate $T(z) = \tau_0 \; \forall z$, this gives rise to a smooth earnings distribution with density $h_0(z)$.

Now introduce a kink, i.e. a change in the marginal tax rate at earnings level z^*, such that the net-of-tax rate decreases from $1-\tau_0$ to $1-\tau_1$. Appendix Figure A5 shows the corresponding budget set and illustrates predicted bunching responses at the kink. One can identify a marginal bunching individual whose indifference curve is tangent to the original budget line at earnings level $z^* + \Delta z^*$ and who is tangent to the upper part of the kinked budget set at z^*. All individuals initially located between z^* and $z^* + \Delta z^*$ bunch. Hence, the bunching mass B^* is given by

$$B^* = \int_{z^*}^{z^* + \Delta z^*} h_0(z) dz \tag{1}$$

where $h_0(z)$ is the counterfactual density of earnings in the absence of the kink.

4.2 Adjustment Frictions: Standard Model with Bunching from Above

Our reduced-form evidence speaks against the frictionless model, which assumes that individuals can costlessly bunch when a kink is introduced at z^* and de-bunch when it is removed. These responses can be attenuated by various types of frictions. Individuals may face costs of adjusting labor supply by renegotiating contracts with employers or searching for new jobs offering the desired hours-wage package, but also costs of paying attention or acquiring information about

---

[11] Following the literature, we define virtual income $y = z - T(z) - [1 - T'(z)]z$ to re-write the nonlinear budget constraint in linearized form.

the tax code. Specifically, we consider a fixed cost of changing earnings between periods, such that utility in period t is given by

$$u(c_t, z_t; n) - \phi \mathbb{1}(z_t \neq z_{t-1})$$

An important factor in modeling bunching responses to a kink under adjustment frictions is given by individuals' initial earnings levels, as these determine the magnitude of the required adjustments.[12] We first consider the standard model following Gelber et al. (2020), where workers are assumed to be initially located at their optimal counterfactual earnings level above the kink. In Section 4.3, we present an alternative framework which takes into account that bunching can originate from below when a kink is created by shifting an existing kink upwards.

4.2.1 Introducing a Kink

When the kink is introduced, frictions attenuate bunching because for some individuals, the utility gain from adjusting their taxable income to z^* is smaller than the adjustment cost they would have to incur. This is illustrated in Panel (a) of Figure 6. An individual with ability $\underline{n}$ initially earns $\underline{z}$ along the budget constraint with tax rate τ_0 (point A). When the kink is introduced, the individual faces a higher tax rate τ_1. She may decide to remain at earnings level $\underline{z}$ (point B), or incur the adjustment cost and reduce her earnings to bunch at z^* (point C). The individual shown in the figure is the "lower" marginal buncher who is indifferent between points B and C. Thus, all individuals with initial earnings between z^* and $\underline{z}$ do not bunch due to adjustment frictions. Individuals with initial earnings between $\underline{z}$ and $z^* + \Delta z^*$ bunch, as the the larger utility difference between their initial earnings level and the kink makes it worthwhile to incur the adjustment cost.[13] Thus, the bunching mass is given by

$$B_1 = \int_{\underline{z}}^{z^*+\Delta z^*} h_0(z)dz \tag{2}$$

Bunching can be further characterized by the indifference condition of the lower marginal buncher

$$u(\underline{z}, \tau_1; \underline{n}) = u(z^*, \tau_1; \underline{n}) - \phi \tag{3}$$

together with the tangency condition of the lower marginal buncher at $\underline{z}$ under τ_0, and the tangency conditions of the upper marginal buncher at $z^* + \Delta z^*$ under τ_0 and at z^* under τ_1.

---

[12] In contrast, individuals' initial earnings and the sequence of tax reforms do not matter for the predicted magnitude of bunching in the frictionless model, where adjustments to the new optimum are instantaneously made from any initial location.

[13] Throughout the analysis, we follow the literature and assume that the benefit of relocating to the kink is increasing in the distance between initial earnings and the kink. See Gelber et al. (2020) for a discussion of this assumption.

4.2.2 Eliminating a Kink

When a kink is eliminated, bunching would immediately disappear in the frictionless model. With adjustment costs, however, some residual bunching remains at the former kink. In Panel (b) of Figure 6, an individual with ability $\overline{n}$ would earn $\overline{z}$ under tax rate τ_0 (point B). When the kink is introduced, the individual strictly prefers bunching at z^*. When the kink is subsequently removed, this individual becomes the marginal de-buncher, as she is indifferent between remaining at the kink (point A) and incurring the adjustment cost to return to her counterfactual earnings level $\overline{z}$. Individuals with initial earnings between $\underline{z}$ and $\overline{z}$ continue to bunch, as the small utility difference between the kink and their counterfactual earnings does not make it worthwhile to de-bunch, but individuals with initial earnings between $\overline{z}$ and $z^* + \Delta z^*$ do not bunch any longer. The remaining bunching mass is

$$B_0 = \int_{\underline{z}}^{\overline{z}} h_0(z)dz \tag{4}$$

Bunching is characterized by the indifference condition of the marginal de-buncher

$$u(z^*, \tau_1; \overline{n}) = u(\overline{z}, \tau_0; \overline{n}) - \phi \tag{5}$$

together with the tangency condition of the marginal de-buncher at $\overline{z}$ under τ_0, and the indifference condition of the lower marginal buncher from equation (3) as well as her tangency condition at $\underline{z}$ under τ_0.

4.3 Adjustment Frictions: Adapted Model with Bunching from Below

In our empirical setting, new kinks in the tax schedule are created by shifting the location of an existing tax bracket cutoff upwards. This situation is typical for income tax settings, where the introduction of kinks from scratch, i.e. from an initial situation without any existing kink, is rarely observed. In this section, we present an adapted framework, taking into account that workers may already have adjusted their earnings in response to a pre-existing kink when analyzing responses to the newly created kink. Note that we abstract from income effects in our main analysis for simplicity, an issue to which we return in Section 4.3.3.

4.3.1 Introducing a Kink

Suppose that the new kink at z^* is introduced while a pre-existing kink at z_0^* of the same size is eliminated, where $z_0^* < z^*$. Bunching at the new kink differs from Section 4.2 because potentially bunching individuals are located at earnings levels different from their counterfactual earnings in the period before the kink is introduced. In fact, if individuals responded fully to the old kink and absent income effects, all potential bunchers at the new kink would be located below z^* under the old kink. Empirical support for the notion of bunching from below is provided by the evidence from Section 3.3. To see why this occurs theoretically, note that the individual with counterfactual earnings $z^* + \Delta z^*$ responds to the change in the net-of-tax rate from $1 - \tau_0$ to $1 - \tau_1$ by reducing earnings by Δz^*. Hence, individuals with counterfactual earnings between z^* and $z^* + \Delta z^*$ are located at or below z^* under the old kink.

In Panel (a) of Figure 7, an individual with ability $\underline{\underline{n}}$ would earn $\underline{\underline{z}}$ under linear tax rate τ_0 (point A). She changed her earnings to $\underline{\underline{z}}'$ in response to the reduction in the net-of-tax rate induced by the old kink. When the new kink is introduced, she may decide to remain at earnings level $\underline{\underline{z}}'$ (point B), or incur the adjustment cost and increase her earnings to bunch at z^* (point C). This individual is the marginal buncher who is indifferent between points B and C.[14] All individuals with counterfactual earnings between z^* and $\underline{\underline{z}}$ bunch, as they moved the furthest below the new kink before it was introduced, and the utility difference between this earnings level and the kink makes it worthwhile incurring the adjustment cost. Individuals with counterfactual earnings between $\underline{\underline{z}}$ and $z^* + \Delta z^*$, who moved less far below z^* under the old kink, do not bunch due to adjustment frictions. Thus, the bunching mass is

$$\tilde{B}_1 = \int_{z^*}^{\underline{\underline{z}}} h_0(z)dz \tag{6}$$

The equation reveals a crucial difference to the standard case described by equation (2). Under the adapted model, bunching originates from below, and the bunchers are those individuals whose counterfactual earnings would be just above the kink. Bunching can be further characterized by the indifference condition of the marginal buncher

$$u(\underline{\underline{z}}', \tau_0; \underline{\underline{n}}) = u(z^*, \tau_1; \underline{\underline{n}}) - \phi \tag{7}$$

together with her tangency condition at $\underline{\underline{z}}$ under τ_0 and her tangency condition at $\underline{\underline{z}}'$ under τ_1.

4.3.2 Eliminating a Kink

Now suppose that the new kink is in turn eliminated. In Panel (b) of Figure 7, an individual with ability $\overline{\overline{n}}$ would earn $\overline{\overline{z}}$ under tax rate τ_0. When the kink is introduced, the individual strictly prefers bunching at z^*. When the kink is subsequently removed, this individual becomes the marginal de-buncher, as she is indifferent between remaining at the kink (point A) and incurring the adjustment cost to return to her counterfactual earnings level $\overline{\overline{z}}$ (point B). Individuals with counterfactual earnings between z^* and $\overline{\overline{z}}$ continue to bunch, but individuals with counterfactual

---

[14] We assume that the distance between z_0^* and z^* is sufficiently large such that the marginal buncher is initially located at a point of tangency above z_0^*. This assumption generally holds in our empirical setting, since the distance between old and new kinks always exceeds implied earnings responses, which we calibrate in Appendix D.2. Furthermore, we note that the analysis here also applies when kinks are shifted over a very large distance, as long as individuals respond to the change in marginal tax rate between kinks.

earnings between $\overline{\overline{z}}$ and $\underline{\underline{z}}$ do not bunch any longer. The remaining bunching mass is

$$\tilde{B}_0 = \int_{z^*}^{\overline{\overline{z}}} h_0(z)dz \tag{8}$$

Bunching is characterized by the indifference condition of the marginal de-buncher

$$u(z^*, \tau_1; \overline{\overline{n}}) = u(\overline{\overline{z}}, \tau_0; \overline{\overline{n}}) - \phi \tag{9}$$

together with her tangency condition at $\overline{\overline{z}}$ under τ_0. Thus, de-bunching occurs towards earnings levels above the kink as in the standard model, and is governed by the same individual conditions in equations (5) and (9). However, the predicted amount and composition of residual bunching differ. In the adapted model, larger bunching remains at the former kink (for given adjustment costs), since there are many bunchers whose counterfactual earnings would be just above z^* and for whom incurring adjustment costs to de-bunch is not worthwhile.

4.3.3 Income Effects

Income effects can somewhat complicate the analysis in the adapted model. In Appendix D.1, we provide the full model solution with income effects. Intuitively, introducing a new kink at z^* not only increases the marginal net-of-tax rate between z_0^* and z^*, but also shifts the budget constraint upwards. This upward shift entails income effects inducing individuals to slightly reduce their earnings. As in the model without income effects, the change in the marginal net-of-tax rate leads to bunching from below at the new kink. Income effects can lead to additional bunching from above, which raises a theoretical possibility of bunching originating from two disjoint segments.

We argue that abstracting from income effects in the main analysis is a useful simplification for two main reasons. First, in Appendix D.2 we present calibration exercises suggesting that income effects are likely of limited empirical relevance in our setting. We calculate the change in virtual income at each of the kinks used in our empirical specifications, and we calibrate potential earnings responses using estimates of the marginal propensity to earn from the literature. We find that income effects are small because (i) virtual income varies only modestly around the tax reforms we study, and (ii) the literature tends to find small earnings responses to changes in unearned income. In particular, the calibration results suggest that any potential bunchers from above would be too close to the kink to make incurring adjustment costs worthwhile, such that bunching from an additional segment due to income effects is unlikely to occur. Thus, in our setting bunching responses are well approximated by the simplified model without income effects. Second, abstracting from income effects is closely in line with the empirical implementation, where we assume quasi-linear utility. This parametrization strategy follows the typical approach in the bunching literature and ensures tractability of the structural bunching estimation.[15]

4.4 Dynamic Models with Adjustment Frictions

Next, we consider a dynamic framework with adjustment frictions, explicitly modeling the trajectory of bunching and de-bunching at kinks over several periods as observed in the reduced-form results. We follow Gelber et al. (2020) and write flow utility in period t as

$$v(c_t, z_t; n, z_{t-1}) = u(c_t, z_t; n) - \tilde{\phi}_t \mathbb{1}(z_t \neq z_{t-1})$$

Adjustment costs $\tilde{\phi}_t$ follow a discrete distribution

$$\tilde{\phi}_t = \begin{cases} \phi & \text{with probability } \pi_{t-t^*} \\ 0 & \text{with probability } 1 - \pi_{t-t^*} \end{cases}$$

where $t - t^*$ is the time since the last policy change. The per-period budget constraint is $z_t - T(z_t) - c_t \geq 0$. Like Gelber et al. (2020), we assume that agents are not forward-looking, leading to a simple decision rule.[16] In each period, individuals move to their optimal frictionless earnings level if the flow utility of moving exceeds the drawn adjustment cost, and remain at their previous earnings level otherwise.

4.4.1 Standard Model with Bunching from Above

First, we set out the dynamic version of the standard model with adjustment frictions, where bunching originates from above. We consider the following timing: In period 0, a linear tax schedule with tax rate τ_0 applies. In period 1, a kink is introduced such that a higher marginal tax rate τ_1 applies at z^* and above. The kink remains in place for T periods. In period $T+1$, the kink is removed.

While the kink is in place in periods $t \in [1, T]$, bunching is

$$\begin{aligned} B_1^t &= \int_{\underline{z}}^{z^*+\Delta z^*} h_0(z)dz + \left(1 - \prod_{j=1}^{t} \pi_j\right) \int_{z^*}^{\underline{z}} h_0(z)dz \\ &= \prod_{j=1}^{t} \pi_j \cdot B_1 + \left(1 - \prod_{j=1}^{t} \pi_j\right) \cdot B^* \end{aligned} \tag{10}$$

As the first line of the equation shows, bunching in period t has two components. The first term

---

[15] To our knowledge, Bastani and Selin (2014) is the only existing study explicitly incorporating income effects into bunching estimation. Their results suggest that the presence of income effects introduces little bias into standard bunching estimates.

[16] This assumption is supported by the lack of anticipatory behavior in the data. Empirically, we do not observe any bunching in the year before kinks are in place, and de-bunching does not begin before kinks are removed.

captures individuals initially located between $\underline{z}$ and $z^* + \Delta z^*$ who bunch immediately regardless of the adjustment cost they draw. The second term corresponds to individuals initially located between the kink and $\underline{z}$ who only bunch once they draw zero adjustment cost. The second line shows that bunching can be expressed as a weighted sum of static bunching with frictions (equation (2)) and frictionless bunching (equation (1)).

When the kink is eliminated in periods $t > T$, bunching is

$$\begin{aligned} B_0^t &= \prod_{j=1}^{t-T} \pi_j \cdot \int_{\underline{z}}^{\overline{z}} h_0(z)dz + \prod_{j=1}^{t-T} \pi_j \cdot \left(1 - \prod_{j=1}^{T} \pi_j\right) \cdot \int_{z^*}^{\underline{z}} h_0(z)dz \\ &= \prod_{j=1}^{t-T} \pi_j \cdot \left[B_0 + \left(1 - \prod_{j=1}^{T} \pi_j\right) \cdot (B^* - B_1)\right] \end{aligned} \tag{11}$$

Hence, residual bunching encompasses two groups of individuals. The first group are those initially located between $\underline{z}$ and $\overline{z}$ who bunched immediately regardless of adjustment costs, and de-bunch once they draw zero costs. The second group are individuals initially located between the kink and $\underline{z}$ who bunched if they drew zero adjustment cost in any period up until T, and who de-bunch once they draw zero costs. Bunching can then be expressed as a weighted sum of static residual bunching (equation (4)) and the difference between bunching with and without frictions (equations (1) and (2)).

4.4.2 Adapted Model with Bunching from Below

Second, we consider a dynamic version of the adapted model, where bunching originates from below. The timing is the following: In period 0, individuals are located at their optimal earnings under the old kink at $z_0^* < z^*$. The new kink at z^* is introduced in period 1 and removed in period $T + 1$.

Bunching at the kink in periods $t \in [1, T]$ is

$$\begin{aligned} \tilde{B}_1^t &= \int_{z^*}^{\underline{\underline{z}}} h_0(z)dz + \left(1 - \prod_{j=1}^{t} \pi_j\right) \int_{\underline{\underline{z}}}^{z^*+\Delta z^*} h_0(z)dz \\ &= \prod_{j=1}^{t} \pi_j \cdot \tilde{B}_1 + \left(1 - \prod_{j=1}^{t} \pi_j\right) \cdot B^* \end{aligned} \tag{12}$$

Again, bunching has two components, individuals who bunch immediately regardless of adjustment cost and individuals who only bunch once they draw zero cost. In line with the static results, the difference to the standard model is that different sets of individuals make up these two groups. Moreover, bunching can be expressed as a weighted sum of static bunching with frictions (equation (6)) and frictionless bunching (equation (1)).

Bunching in periods $t > T$ is

$$\tilde{B}_0^t = \prod_{j=1}^{t-T} \pi_j \cdot \int_{z^*}^{\overline{\overline{z}}} h_0(z)dz = \prod_{j=1}^{t-T} \pi_j \cdot \tilde{B}_0 \tag{13}$$

Thus, dynamic residual bunching consists of individuals initially located between z^* and $\overline{\overline{z}}$, corresponding to a subset of residual bunchers from equation (8). In the dynamic case, these individuals bunched immediately regardless of the adjustment cost they drew, and de-bunch once they draw zero adjustment costs.[17]

4.5 Empirical Implementation

Given the available variation in tax rates and the bunching moments from the reduced-form analysis, it is possible to estimate the parameters of the different model versions by specifying a functional form of utility $u(c, z; n)$. Following the typical approach in the bunching literature, we adopt iso-elastic, quasi-linear preferences:

$$u(c, z; n) = c(z) - \frac{n}{1 + \frac{1}{\varepsilon}} \left(\frac{z}{n}\right)^{1+\frac{1}{\varepsilon}}$$

where ε is the structural elasticity of taxable income w.r.t. the net-of-tax rate. Under these preferences, optimal earnings supply at tax rate τ is then given by $z = n(1 - \tau)^{\varepsilon}$.

We can use the static models to estimate the two parameters ε and ϕ. Equations (3) and (5), and equations (7) and (9), respectively, describe the relationship between predicted bunching and these parameters in the standard model and the adapted model. In the dynamic models, predicted bunching in a period is a function of the probability of drawing positive adjustment costs in addition to ε and ϕ (equations (10) and (11); (12) and (13)). In the data, we observe kinks remaining in place for up to three years, and thus the cumulative probabilities π_1, $\pi_1\pi_2$ and $\pi_1\pi_2\pi_3$ can be estimated by the dynamic models. Appendix C shows the detailed estimation equations of all model versions.

Moreover, we allow for adjustment frictions to be asymmetric in some specifications. This captures the notion that the cost of adjusting earnings precisely towards a kink may differ from the cost of adjusting earnings away from the kink. In this case, we replace the single adjustment cost parameter ϕ by a bunching adjustment cost parameter ϕ_+ in the bunching equations (3) and (7), and by a de-bunching adjustment cost parameter ϕ_- in the de-bunching equations (5) and (9). We can then estimate ε, ϕ_+ and ϕ_- based on the static model versions, and these parameters plus π_1, $\pi_1\pi_2$ and $\pi_1\pi_2\pi_3$ based on the dynamic models.

---

[17] Two further groups of initial bunchers do not generate any residual bunching. First, individuals initially located between $\overline{\overline{z}}$ and $\underline{\underline{z}}$ also bunched immediately regardless of the adjustment costs they drew up until period T, but de-bunch immediately in period $T + 1$. Second, individuals initially located between $\underline{\underline{z}}$ and $z^* + \Delta z^*$ bunched if they draw zero adjustment costs up until period T, but also de-bunch immediately in period $T + 1$.

As empirical moments, we use observed bunching while kinks are in place and residual bunching after they are removed. Our preferred specifications combine the information from the full set of kinks from Figure 3 in all available years.[18] In addition, we present results based on a narrower set of moments, using only the CYP 10,000 kink in the years 2004 to 2008, where we observe the full trajectory of bunching and de-bunching at the same kink. We use a minimum distance estimator to estimate parameters in all specifications, matching predicted bunching and the bunching observed in the data as closely as possible. To obtain standard errors, we bootstrap the entire estimation procedure, re-estimating the structural parameters for each set of bootstrapped bunching estimates.

5 Estimation Results

Table 2 presents the main structural estimation results. In Panel A, we begin with results based on estimating the standard model, where bunching is assumed to originate from above. Using only the CYP 10,000 kink in the years 2004 to 2008, the model yields a sizeable adjustment cost of CYP 54 and a taxable income elasticity of 0.04. Our preferred specification based on the full set of kinks yields a larger adjustment cost of CYP 140 and an elasticity of 0.22. Next, Panel B shows results from a specification allowing for asymmetric frictions. The estimates suggest important asymmetry in adjustment costs. Bunching and de-bunching at the CYP 10,000 kink reveals a bunching adjustment cost of CYP 128 but a de-bunching adjustment cost of only CYP 22. Similarly, when exploiting the full set of kinks, bunching costs are CYP 127 and de-bunching costs are CYP 44.[19] The elasticity estimates vary between 0.04 and 0.07.

Panel C of the table shows analogous estimates based on the adapted model with bunching from below. The evidence of asymmetry in adjustment costs is even starker. Bunching costs are estimated at CYP 63 at the CYP 10,000 kink and CYP 49 across the full set of kinks, whereas the estimation yields very small de-bunching costs of at most CYP 2.[20] The adapted model also implies a somewhat higher taxable income elasticity of 0.25 to 0.29. Interestingly, it appears that allowing for asymmetric frictions is crucial in order to obtain meaningful adjustment cost estimates under the adapted model. Appendix Table A3 shows extended structural estimation results, including from a version of the adapted model with symmetric frictions. When adjustment frictions are not allowed to vary between bunching and de-bunching, the adapted model

---

[18] As discussed in Section 3.2, there are other kinks, in particular those at higher earnings levels, at which we do not observe significant bunching. We do not use these moments in the structural estimation, since the absence of bunching could be driven either by a small elasticity or large adjustment costs around these kinks and the model cannot separately identify these two reasons. In this sense, our estimates should be interpreted as local to the earnings levels around the kinks we include.

[19] T-tests confirm that the difference between bunching and de-bunching adjustment costs is statistically significant at the 5% level in all specifications shown in Table 2.

[20] Note that we impose a lower bound of CYP 1 on adjustment costs in the estimation. This restriction is binding only in some specifications based on the adapted model. Whenever the estimation algorithm selects the lower bound, we conclude that adjustment costs are at most CYP 1 and thus not economically meaningful.

yields very low symmetric adjustment cost estimates of at most CYP 7. Furthermore, Appendix Table A3 includes additional results based on an intermediate set of kinks, excluding the kinks at CYP 9,000 and CYP 10,750, where the reduced-form evidence of adjustment frictions is less clear. Parameter estimates from this alternative set of moments tend to be similar to those based on the full set of kinks, if anything yielding somewhat larger adjustment costs in most specifications.

Table 3 shows results from estimating the dynamic model variants. In Panel A, the dynamic version of the standard model yields large adjustment cost estimates of CYP 98 at the CYP 10,000 kink and CYP 144 across the full set of kinks. Elasticity estimates are 0.04 and 0.24, respectively. The cumulative probabilities of facing adjustment frictions are 0.62 to 0.65 in the first year after the policy change, 0.06 to 0.28 in the second year, and zero to 0.02 in the third year. The dynamic version of the adapted model in Panel B yields smaller adjustment costs between CYP 11 and CYP 14 and an elasticity of 0.10 to 0.13. The adapted model also implies smaller probabilities of facing adjustment frictions of 0.03 to 0.08 in the first year, and close to zero already in the second year. Taken together, the results suggest that frictions dissipate after two to three years.[21] Finally, Appendix Table A4 presents an extended set of dynamic estimates, including results from dynamic models with asymmetric frictions and results based on an alternative set of bunching moments. Overall, the estimated parameters are of similar magnitude to the main table, and the dynamic models again suggest clear asymmetry in adjustment frictions. For instance, the standard model yields a bunching cost of CYP 129 and a de-bunching cost of only CYP 17 across the full of kinks.

Goodness of Fit. Appendix Table A5 compares the fit of the different models we estimate. In general, the estimation closely matches predicted and observed bunching, leading to small estimation errors. Among static models, the standard model with asymmetric frictions provides the best fit both at the CYP 10,000 kink and among all kinks. The adapted model performs substantially better than the standard model with symmetric frictions when using all kinks as estimation moments, but slightly worse when using only the CYP 10,000 kink. Out of the dynamic specifications, the adapted model generally provides the better fit. Overall, goodness of fit improves when allowing for asymmetries and dynamics, but it does not seem to clearly favor either the standard or the adapted model.

Discussion. While the exact magnitude of parameter estimates remains somewhat sensitive to the model used, we take away several useful insights from the structural estimation. First, we find economically meaningful adjustment frictions across our preferred specifications. For

---

[21] We observe kinks in place for up to three years and thus we cannot explicitly estimate the probability of facing frictions beyond this time horizon. Yet, since the estimated cumulative probability is already close to zero in the third year in most specifications, our results imply very small upper bounds on the cumulative probabilities in the fourth year and later. The finding that frictions dissipate after two to three years also closely mirrors results from the literature (e.g. Gelber et al. 2020).

instance, the results from Table 2 imply that taxpayers around the kink would have to incur costs representing around 0.5% to 1.5% of their taxable income to adjust. These costs are especially substantial compared to their tax liability which is close to zero at this level of taxable income. Second, our findings indicate an important asymmetry in the cost of adjusting earnings, where bunching is costlier than de-bunching. A natural interpretation of this asymmetry may be that adjusting taxable income precisely towards the earnings level at a kink entails greater difficulty than adjusting away from the kink, where such precision is not required. Earnings may also naturally drift upwards over time, which may make it relatively easy to de-bunch by increasing earnings. Third, we note that our results on adjustment costs are quite robust with respect to the set of bunching moments used in the estimation. The level of estimated adjustment costs is similar at the CYP 10,000 kink and at all kinks in most specifications, and we find clear asymmetry in bunching vs. de-bunching costs based on either set of moments. Fourth, differences in the estimated structural elasticities and observed elasticities based on the Saez (2010) method are worth considering. At the CYP 10,000 kink, the structural elasticity estimates are as low as 0.04 in some specifications, which is similar to the observed elasticity at this kink in 2006. Taking this comparison at face value would imply that observed bunching at CYP 10,000 may be close to "complete" in this year. Across the full set of kinks, structural elasticity estimates tend to be higher, suggesting that there would be more bunching in a frictionless world than observed in the reduced-form evidence.

Finally, the adapted model is arguably more conceptually appealing given our institutional setting and the evidence that the majority of bunchers come from below, but it does not seem perfectly suited for all estimation purposes. A particular issue with the adapted model lies in de-bunching frictions, which are estimated close to zero, and which can spill over into very small overall adjustment cost estimates when frictions are assumed to be symmetric. What is behind this issue? In the adapted model, the predicted bunchers in equation (6) are individuals whose counterfactual earnings are just above the kink. For these individuals, it is hardly worthwhile to de-bunch even at very low adjustment costs, since the difference between earnings at the kink and their optimal counterfactual earnings is small. Thus, the amount of residual bunching we observe empirically, which is around one half to one third of bunching while the kink in place, can be rationalized by very low adjustment costs in the adapted model. However, this is somewhat inconsistent with the evidence from Figure 4, which suggests that de-bunchers typically move quite far above the kink.

6 Bunching and Adjustment Frictions among the Self-Employed

Our analysis so far concentrates on wage earners. In this section, we study bunching responses and adjustment frictions among self-employed workers, who make up 8.7% of individuals in the data. Interestingly, we find that bunching responses of the self-employed are less pronounced at the kinks at lower earnings levels used in the main analysis. For instance, Appendix Figure A6 shows that in contrast to wage earners, the self-employed do not exhibit significant bunching at the CYP 10,000 kink. There are some small spikes during some of the years between 2004 and 2006, but they are attributed to round-number effects by the estimated counterfactual and thus the excess mass is not significantly different from zero in these years.

However, strong bunching responses of the self-employed can be found at the CYP 20,000 kink. Similarly to the CYP 10,000 kink, this kink is introduced in 2004, remains in place until 2006 and is subsequently removed in 2007, such that we can track the full trajectory of responses. In Figure 8, there is clearly visible bunching at the CYP 20,000 beginning in 2004 (Panel b), when the excess mass is already 3.5. The excess mass then grows steadily to 6.5 in 2005 (Panel c) and 8.3 in 2006 (Panel d). In line with growing bunching, the observed elasticity increases from 0.13 in 2004 to 0.31 in 2006. When the kink is removed in 2007, there is still significant residual excess mass of around 2.7, but bunching finally disappears in 2008.

These bunching dynamics among the self-employed resemble those of wage earners around the CYP 10,000 kink. Again, the gradual growth in excess mass between 2004 and 2006 implies that some workers face frictions in bunching adjustments, and the presence of residual bunching in 2007 suggests frictions in de-bunching adjustments. It is worth noting some important differences, however. First, both excess mass and observed elasticities are larger among the self-employed. The observed elasticity of 0.31 in 2006 of the self-employed is several times larger than the maximum value of 0.07 among wage earners. Second, there is no visible diffuse bunching of self-employed around the kink, implying that they are able to target the kink more precisely than wage earners. Similar differences between wage earners and self-employed have been documented in the literature (e.g. Chetty et al. 2011; Kleven and Waseem 2013; Bastani and Selin 2014). These patterns are consistent with the self-employed facing adjustment frictions to a lesser extent, for instance because they have more autonomy over adjusting their labor supply and earnings. Furthermore, while wage earners tend to bunch only at the first tax kink, the self-employed bunch more strongly at the highest kink. Such differences could be due to heterogeneity in structural elasticities and adjustment costs, and we note that our results for the self-employed are local to higher taxable income levels. Another potential explanation for the lack of significant bunching among the self-employed at the first kink could be the generally smaller sample size, although this may be less likely since we do find significant excess mass at the CYP 20,000 kink.

Finally, we repeat the structural estimation for the self-employed at the CYP 20,000 kink, choosing specifications analogous to those estimated for wage earners from Table 2. Table 4 presents the resulting parameter estimates. In line with the reduced-form evidence, we find generally lower adjustment costs and larger elasticities among the self-employed. Across specifications, adjustment costs are at most CYP 22. When allowing for asymmetric frictions, the bunching adjustment cost is estimated between CYP 7 and CYP 16, and the de-bunching adjustment cost is at most CYP 5. These magnitudes are considerably lower than corresponding estimates for wage earners from Table 2, where symmetric adjustment costs are at least CYP 54 and bunching adjustment costs are at least CYP 49. Furthermore, the estimation yields large structural elasticities between 0.30 and 0.55 for the self-employed.[22]

7 Conclusion

In this paper, we study adjustment costs in responses to income taxes, exploiting a series of tax reforms in Cyprus which create and subsequently eliminate kinks in the tax schedule. Reduced-form evidence reveals significant adjustment frictions in earnings responses to marginal tax rates. Bunching is initially attenuated by frictions and builds gradually over time, and de-bunching does not occur immediately. Employing a structural model of frictional earnings supply, we find sizeable adjustment costs between CYP 54 (EUR 93) and CYP 140 (EUR 238) for wage earners in our preferred specifications. Moreover, our results suggest important asymmetries in adjustment frictions, where bunching exactly at a kink entails greater cost than de-bunching away from the kink. Finally, we estimate considerably lower adjustment costs of at most CYP 22 (EUR 38) among self-employed workers.

These findings have implications for the interpretation of observed responses to income taxes. The sizeable adjustment costs we estimate imply that earnings responses can be substantially attenuated in the short-term. In contrast to most existing studies, we find significant bunching responses among wage earners at various kinks. However, we also find significant adjustment frictions for this group, which may explain the lack of bunching at other kinks and in other settings. Moreover, our results can have valuable policy implications. When adjustment frictions inhibit workers' short-run adjustments to tax reforms, this may be important to take into account in government projections regarding reactions to tax policy. Finally, frequent tax reforms such as those in Cyprus in the 2000s induce individuals to make costly adjustments, which may entail additional welfare costs not accounted for in models without frictions.

---

[22] To test whether parameters among the self-employed are significantly different from those of wage earners, we perform t-tests comparing the estimates from Table 4 to analogous specifications from Table 2. The differences in adjustment costs are significant at the 5% level, except when comparing estimates based on the standard model with symmetric frictions shown in the first lines of the tables. Differences in elasticities are always highly significant.

References

Akee, R., Copland, W., Keeler, G., Angold, A. and Costello, E. (2010), 'Parents' Incomes and Children's Outcomes: A Quasi-Experiment Using Transfer Payments from Casino Profits', *American Economic Journal: Applied Economics* **2**(1), 86–115.

Ashenfelter, O. and Plant, M. (1990), 'Nonparametric Estimates of the Labor-Supply Effects of Negative Income Tax Programs', *Journal of Labor Economics* **8**(1), 396–415.

Bastani, S. and Selin, H. (2014), 'Bunching and Non-bunching at Kink Points of the Swedish Tax Schedule', *Journal of Public Economics* **109**, 36–49.

Blomquist, S., Newey, W., Kumar, A. and Liang, C.-Y. (2021), 'On Bunching and Identification of the Taxable Income Elasticity', *Journal of Political Economy* **129**(8), 2320–2343.

Blomquist, S. and Selin, H. (2010), 'Hourly Wage Rate and Taxable Labor Responsiveness to Changes in Marginal Tax Rates', *Journal of Public Economics* **94**(11–12), 878–889.

Burtless, G. (1986), 'The Work Response to a Guaranteed Income: A Survey of Experimental Evidence', *Conference Series Proceedings, Federal Reserve Bank of Boston* **30**, 22–59.

Cesarini, D., Lindqvist, E., Notowidigdo, M. and Östling, R. (2017), 'The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries', *American Economic Review* **107**(12), 3917–3946.

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

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

Chetty, R., Guren, A., Manoli, D. and Weber, A. (2013), 'Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities', *NBER Macroeconomics Annual 2012* **27**, 1–56.

Farhi, E. and Gabaix, X. (2020), 'Optimal Taxation with Behavioral Agents', *American Economic Review* **110**(1), 298–336.

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* **103**(3), 551–572.

Gelber, A. (2012), 'Taxation and Earnings of Husbands and Wives: Evidence from Sweden', *Review of Economics and Statistics* **96**(2), 287–305.

Gelber, A., Isen, A. and Song, J. (2017), The Role of Social Security Benefits in the Initial Increase of Older Women's Employment: Evidence from the Social Security Notch, *in* C. Goldin and L. Katz, eds, 'Women Working Longer: Increased Employment at Older Ages', pp. 239–268.

Gelber, A., Jones, D. and Sacks, D. (2020), 'Estimating Earnings Adjustment Frictions: Method and Evidence from the Earnings Test', *American Economic Journal: Economic Policy* **12**(1), 1–31.

Giupponi, G. (2019), 'When income effects are large: Labor supply responses and the value of welfare transfers'. Working paper.

Goolsbee, A. (2000), 'What Happens When You Tax the Rich? Evidence from Executive Compensation', *Journal of Political Economy* **108**(2), 352–378.

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

Gudgeon, M. and Trenkle, S. (2021), 'The Speed of Earnings Responses to Taxation and the Role of Firm Labor Demand'. Working Paper.

He, D., Peng, L. and Wang, X. (2021), 'Understanding the Elasticity of Taxable Income: A Tale of Two Approaches', *Journal of Public Economics* **197**. Article 104375.

Imbens, G., Rubin, D. and Sacerdote, B. (2001), 'Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players', *American Economic Review* **91**(4), 778–794.

Jones, D. and Marinescu, I. (2019), 'The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund', *forthcoming, American Economic Journal: Applied Economics* .

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

Kleven, H. and Schultz, E. (2014), 'Estimating Taxable Income Responses Using Danish Tax Reforms', *American Economic Journal: Economic Policy* **6**(4), 271–301.

Kleven, H. and Waseem, M. (2013), 'Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan', *Quarterly Journal of Economics* **128**, 669–723.

Mavrokonstantis, P. (2019), 'Bunching: Estimate bunching'. R package version 0.8.4, `https://CRAN.R-project.org/package=bunching`.

Pashardes, P. and Polycarpou, A. (2010), ‘Labor Supply in Cyprus’. Economic Analysis Papers 05-10, Economics Research Centre, University of Cyprus.

Pashardes, P. and Polycarpou, A. (2012), ‘The Labor Market Behavior of Public Assistance Recipients in Cyprus’. Economic Analysis Papers 09-12, Economics Research Centre, University of Cyprus.

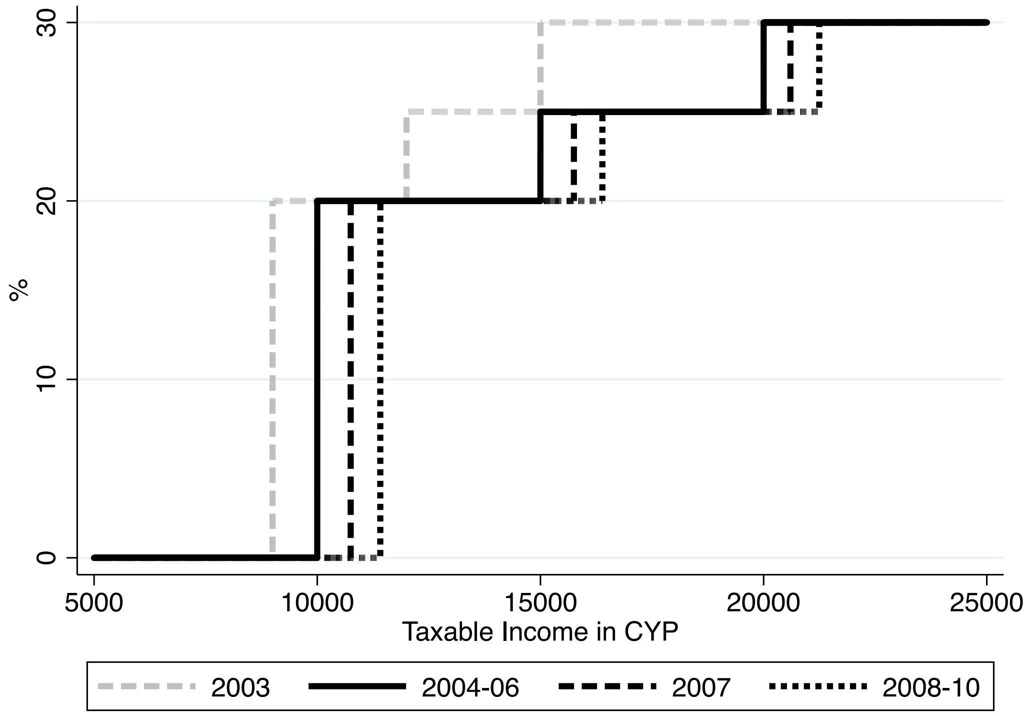
Saez, E. (2010), ‘Do Taxpayers Bunch at Kink Points’, *American Economic Journal: Economic Policy* **2**(3), 180–212.

Saez, E., Slemrod, J. and Giertz, S. (2012), ‘The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review’, *Journal of Economic Literature* **50**(1), 3–50.

Werquin, N. (2016), ‘Income Taxation with Frictional Labor Supply’. Working Paper.

Zaresani, A. (2020), ‘Adjustment Costs and Incentives to Work: Evidence from a Disability Insurance Program’, *Journal of Public Economics* **188**. Article 104223.

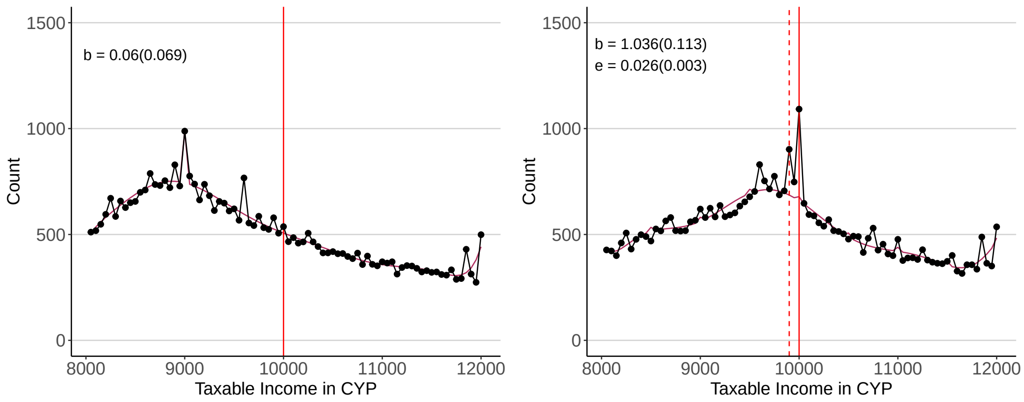
Figure 1: Marginal Tax Rates, 2003 to 2010



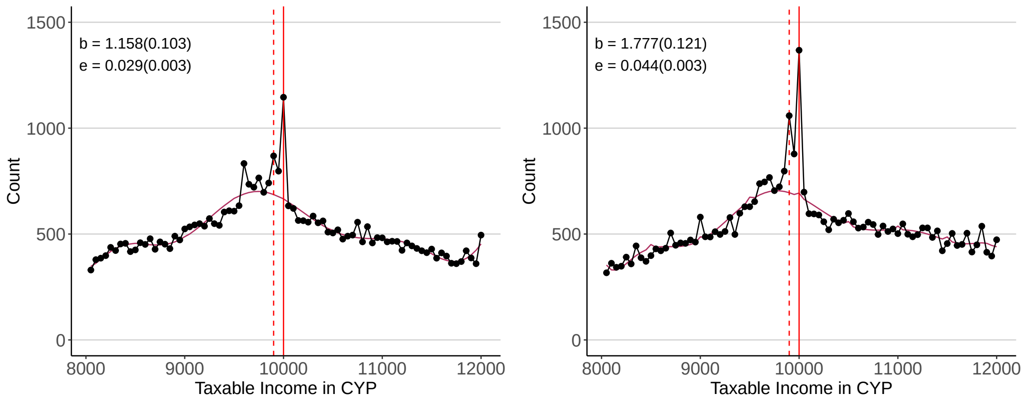
Notes: The figure shows marginal income tax rates by taxable income in the years 2003 to 2010.

Figure 2: Bunching at the CYP 10,000 Kink

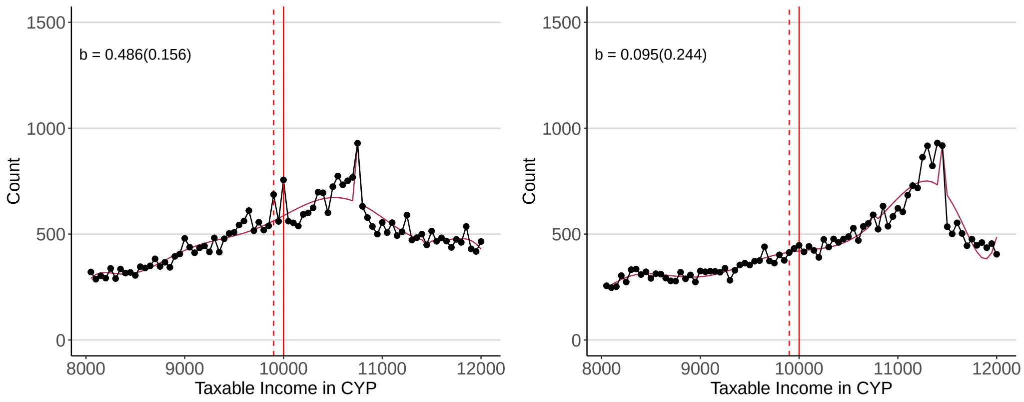
(a) 2003 **(b)** 2004



(c) 2005 **(d)** 2006



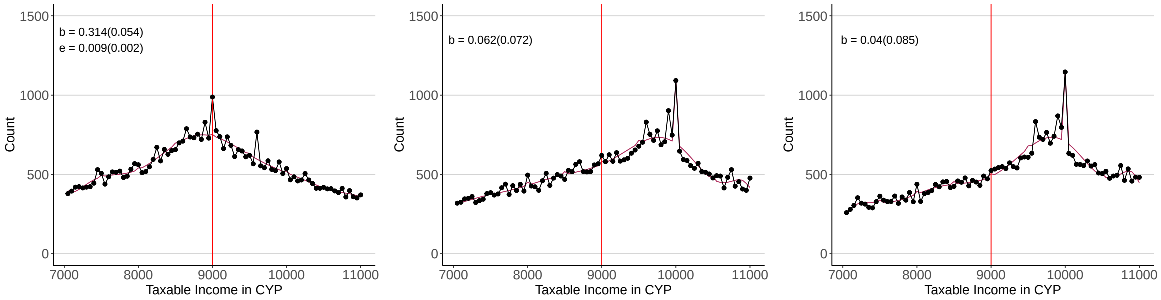
(e) 2007 **(f)** 2008



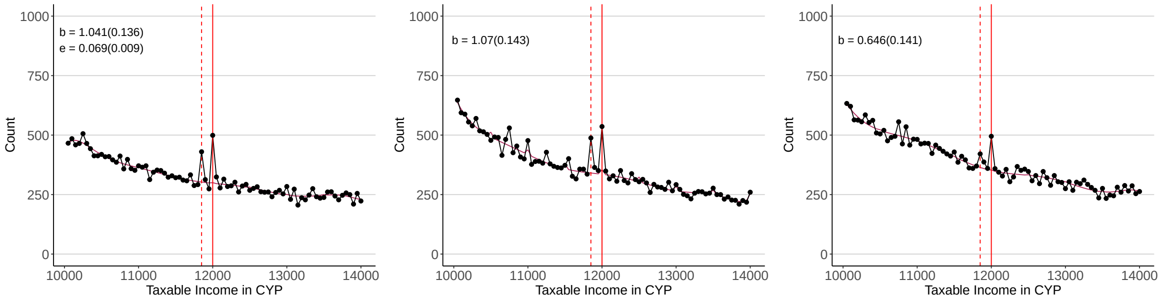
Notes: The figure shows the bunching dynamics of wage earners around the CYP 10,000 kink between 2003 and 2008. The kink is introduced in 2004, remains in place until 2006 and is removed in 2007. In each panel, the black connected dots plot the empirical taxable income distribution in bins of width CYP 50 and the red line is the estimated counterfactual density. Solid vertical lines mark the location of the kink, while dashed lines demarcate lower and upper bounds of the bunching region. The panels also show the estimated excess mass b, and for years 2004 to 2006, the observed elasticity e. Bootstrapped standard errors are in parentheses.

Figure 3: Bunching at Other Tax Kinks

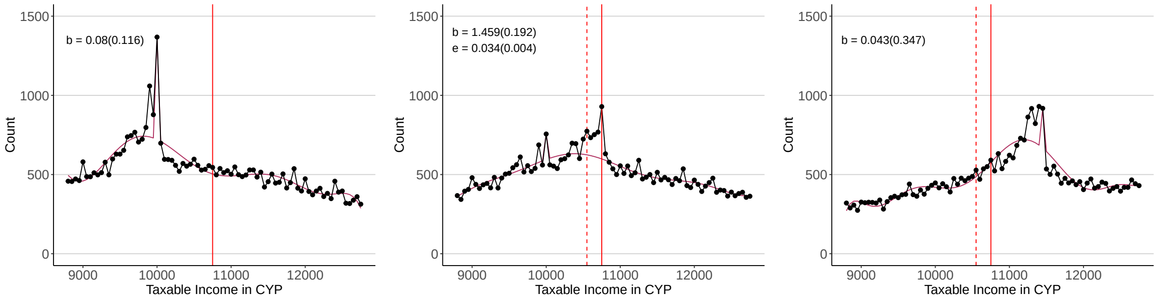
(a) CYP 9,000 kink, 2003 **(b)** CYP 9,000 kink, 2004 **(c)** CYP 9,000 kink, 2005



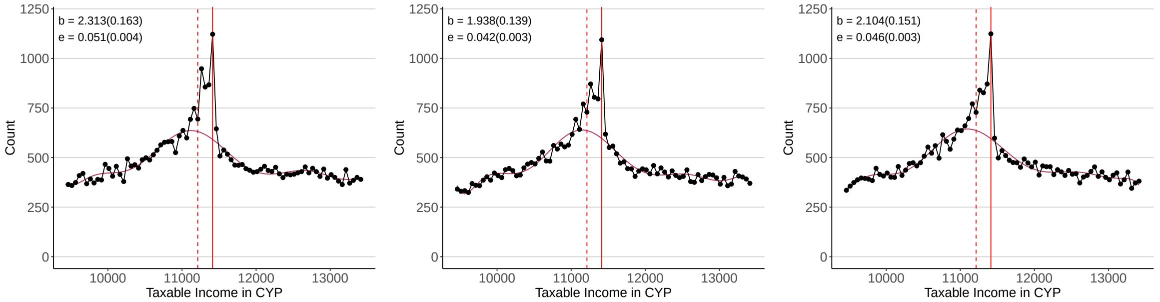
(d) CYP 12,000 kink, 2003 **(e)** CYP 12,000 kink, 2004 **(f)** CYP 12,000 kink, 2005



(g) CYP 10,750 kink, 2006 **(h)** CYP 10,750 kink, 2007 **(i)** CYP 10,750 kink, 2008

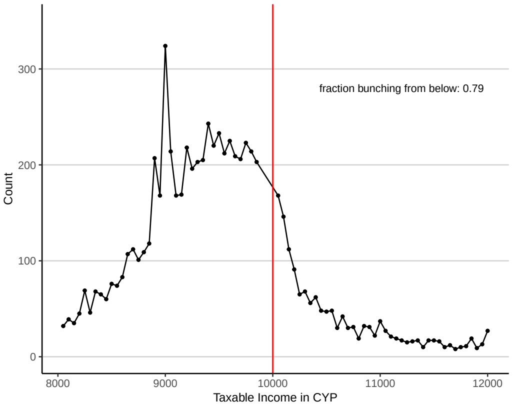


(j) CYP 11,413 kink, 2008 **(k)** CYP 11,413 kink, 2009 **(l)** CYP 11,413 kink, 2010

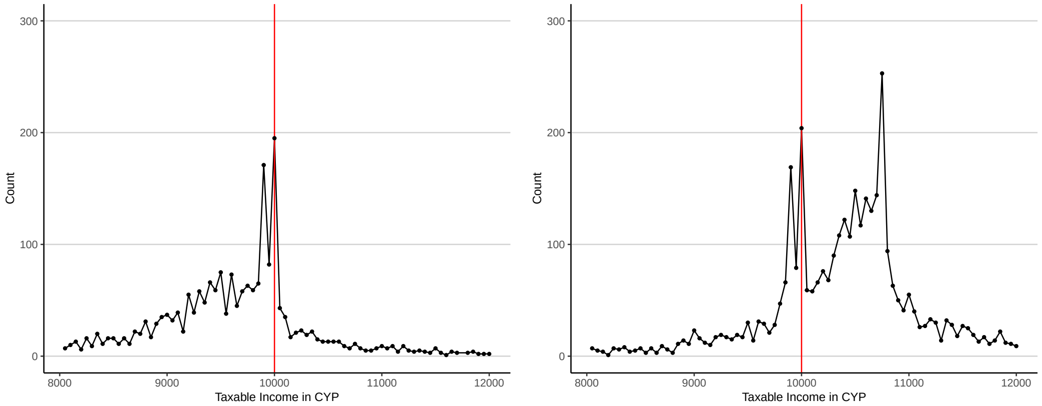


Notes: The figure shows bunching dynamics around different tax kinks. The kinks at CYP 9,000 (Panels a to c) and CYP 12,000 (Panels d to f) are in place in 2003 and removed in 2004. The kink at CYP 10,750 (Panels g to i) is introduced in 2007 and removed in 2008. The kink at CYP 11,413 (Panels j to l) is introduced in 2008 and remains in place until 2010. In each panel, the black connected dots plot the empirical distribution in bins of width CYP 50 and the red line is the estimated counterfactual density. Solid vertical lines mark the location of the kink, while dashed lines demarcate the lower bound of the bunching region. The panels also show the estimated excess mass *b*, and for the years when the respective kink is in place, the observed elasticity *e*. Bootstrapped standard errors are in parentheses.

Figure 4: Dynamics of Individual Bunching and De-Bunching at the CYP 10,000 Kink



(a) Year $t-1$ Income of New Bunchers in 2004 to 2006



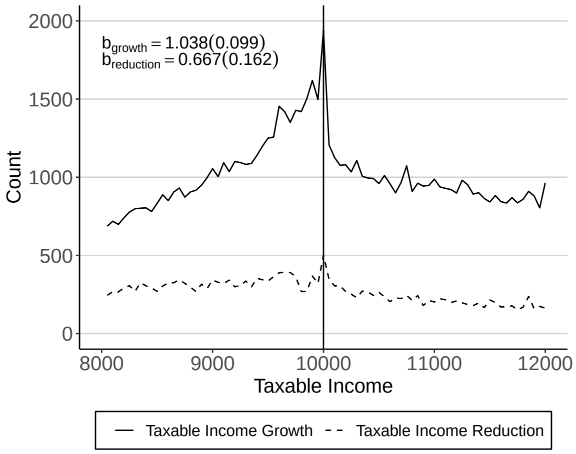
(b) Year $t-1$ Income of Residual Bunchers in 2007

(c) Year $t+1$ Income of Bunchers in 2006

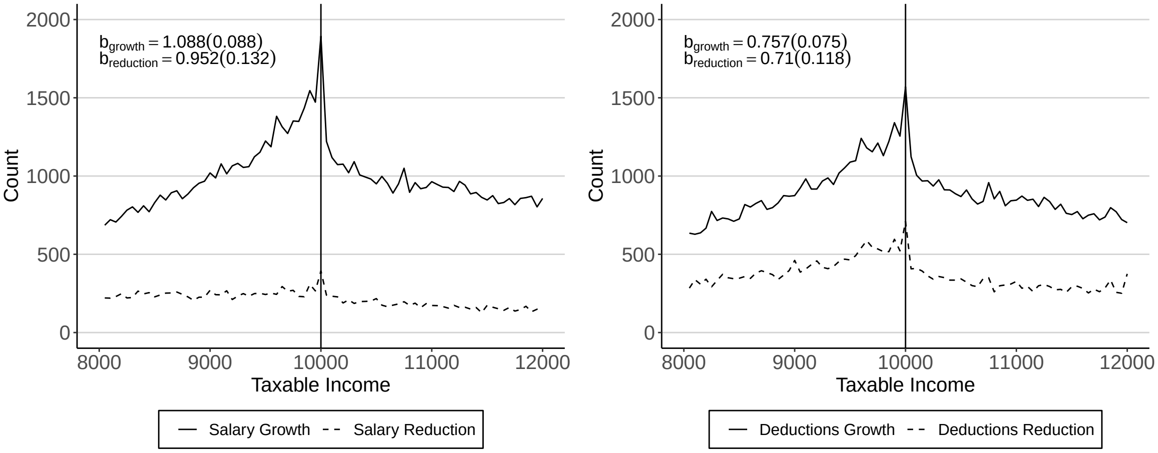
Notes: The figure shows the dynamics of individual bunching and de-bunching at the CYP 10,000 kink. Panel (a) focuses on individuals who bunch at the CYP 10,000 kink in any of the years 2004 to 2006 while it is in place, and shows their distribution of taxable income in the year before they bunch at the kink for the first time. "Fraction bunching from below" denotes the fraction of new bunchers whose taxable income was below the kink in the year before bunching. Panel (b) focuses on individuals who still bunch in 2007 after the kink is removed, and shows their distribution of taxable income in the previous year 2006. Panel (c) focuses on individuals who bunch at the kink in 2006, the last year it is in place, and shows their distribution of taxable income in the next year 2007. In each panel, the black connected dots plot the empirical distribution in bins of width CYP 50 by year. Solid vertical lines mark the location of the CYP 10,000 kink.

Figure 5: Correlation of Bunching with Taxable Income, Gross Earnings and Deductions Growth

(a) Taxable Income Growth



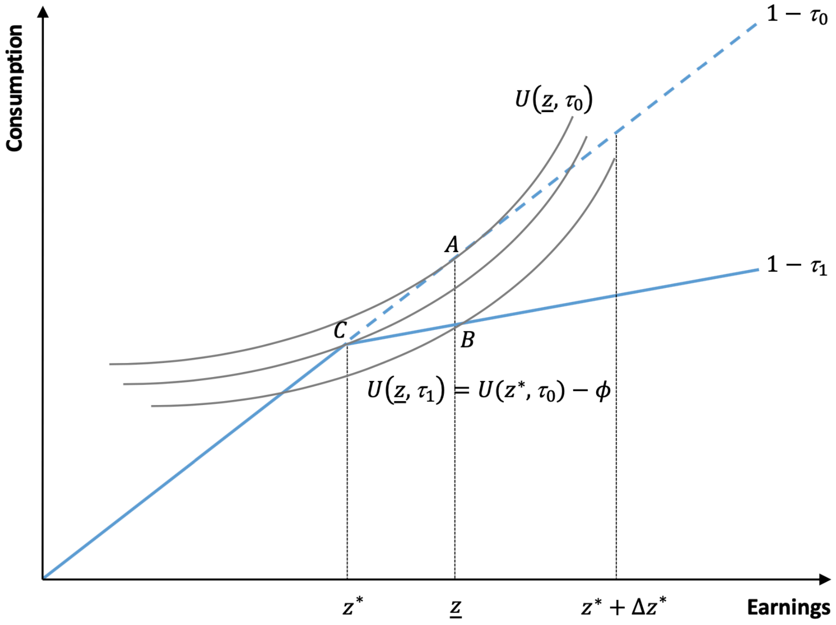
(b) Gross Earnings Growth **(c)** Deductions Growth



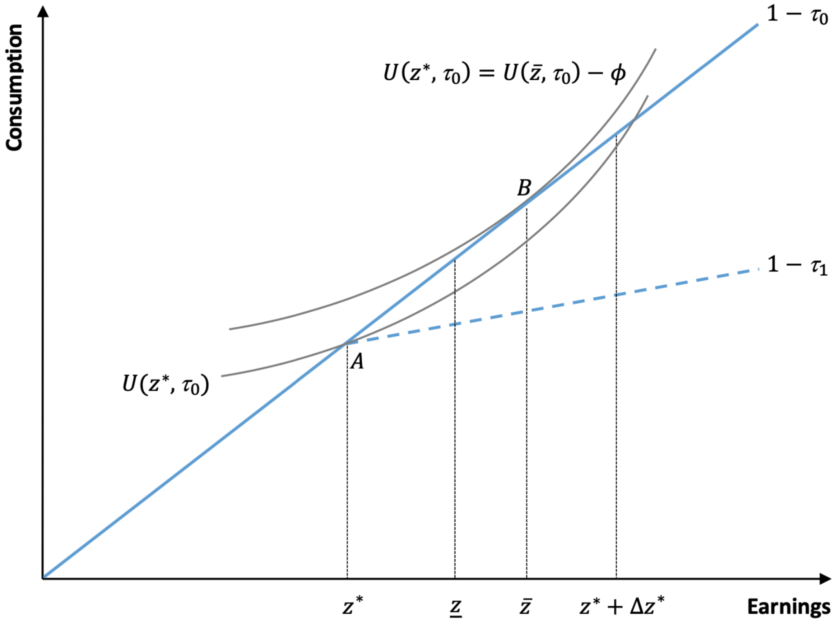
Notes: The figure shows bunching patterns around the CYP 10,000 kink, pooling across the years 2004 to 2006 while it is in place. Each panel shows the distribution of taxable income for subgroups defined by experiencing growth vs. reductions in taxable income (Panel a), gross earnings (Panel b) and deductions (Panel c) relative to the prior year. Each panel also reports, for each sub-sample, the estimated excess mass b with bootstrapped standard errors in parentheses.

Figure 6: Frictions and Bunching from Above in the Standard Model

(a) Frictions Attenuating Bunching



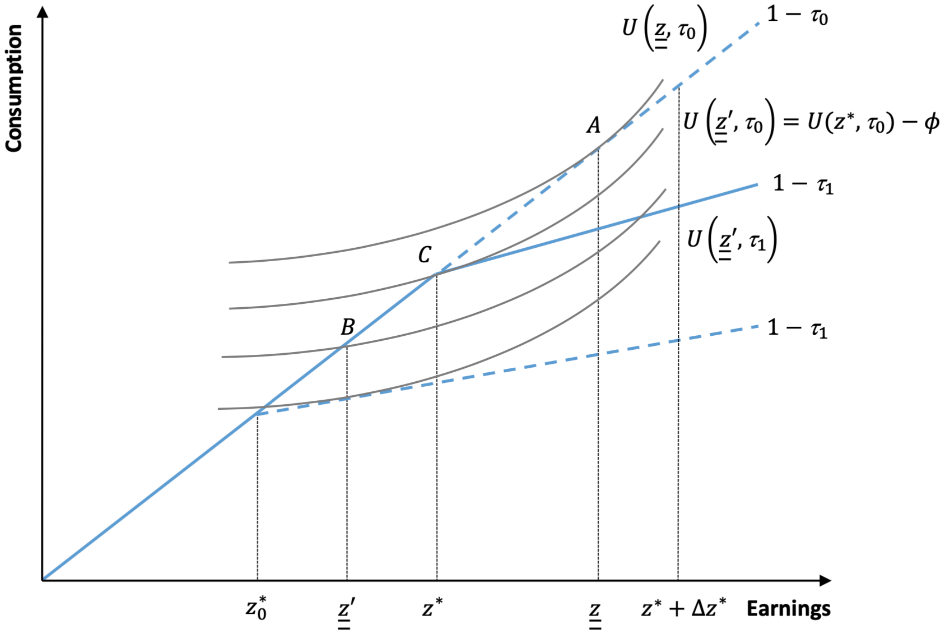
(b) Frictions Attenuating De-Bunching



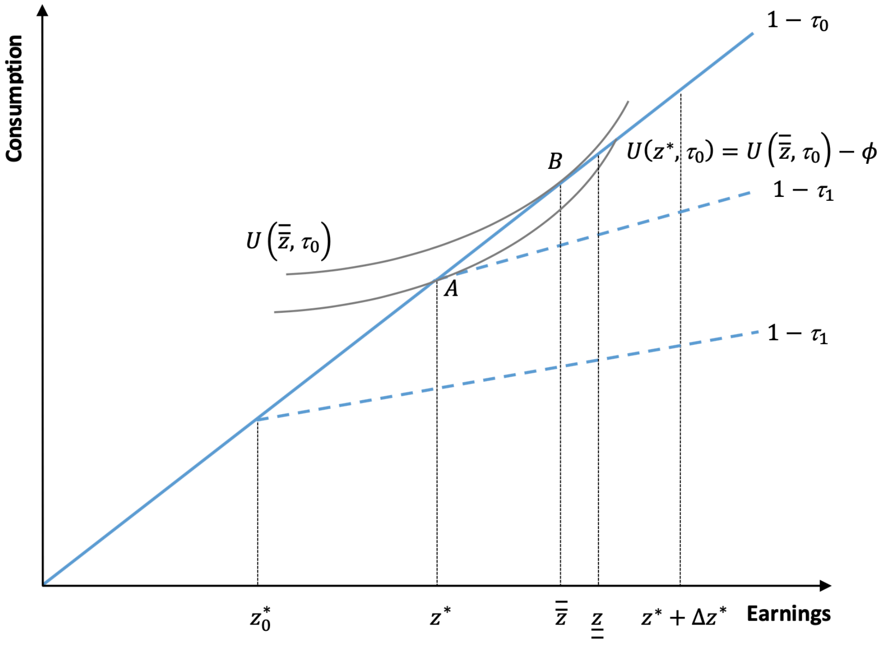
Notes: The figure shows how frictions attenuate bunching responses in the standard model from Section 4.2. Panel (a) shows how bunching is attenuated when a kink is in place. Indifference curves shown are those of the marginal buncher who is indifferent between remaining at her counterfactual earnings $\underline{z}$ and bunching at the kink z^*. Panel (b) shows how de-bunching is attenuated when a kink is removed. The indifference curves are those of the marginal de-buncher who is indifferent between remaining at z^* and de-bunching towards her counterfactual earnings $\overline{z}$.

Figure 7: Frictions and Bunching from Below in the Adapted Model

(a) Frictions Attenuating Bunching



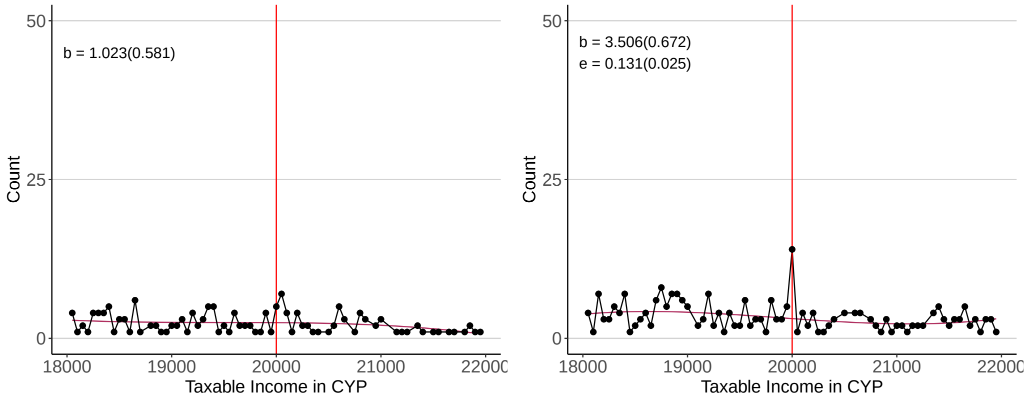
(b) Frictions Attenuating De-Bunching



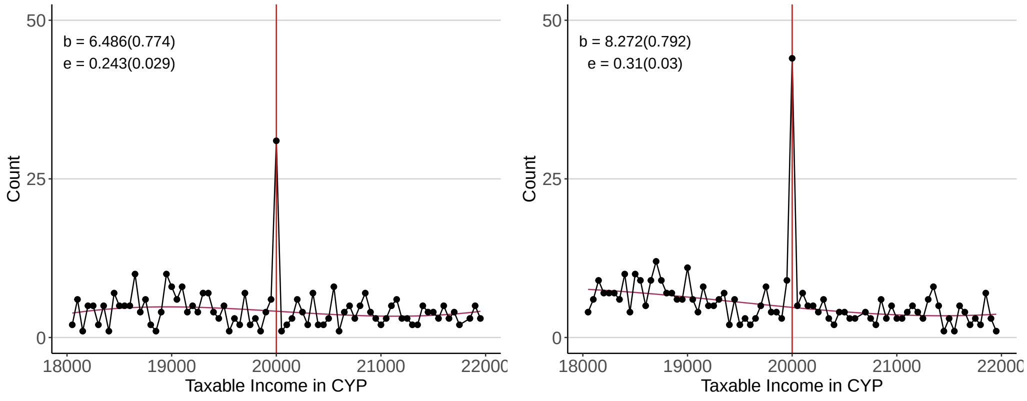
Notes: The figure shows how frictions attenuate bunching responses in the adapted model from Section 4.3. Panel (a) shows how bunching is attenuated when a kink is in place. Indifference curves shown are those of the marginal buncher with counterfactual earnings $\underline{z}$ who is indifferent between remaining at $\underline{z}'$ and bunching at the kink z^*. Panel (b) shows how de-bunching is attenuated when a kink is removed. The indifference curves are those of the marginal de-buncher who is indifferent between remaining at z^* and de-bunching towards her counterfactual earnings $\bar{\bar{z}}$.

Figure 8: Bunching of the Self-Employed at the CYP 20,000 Kink

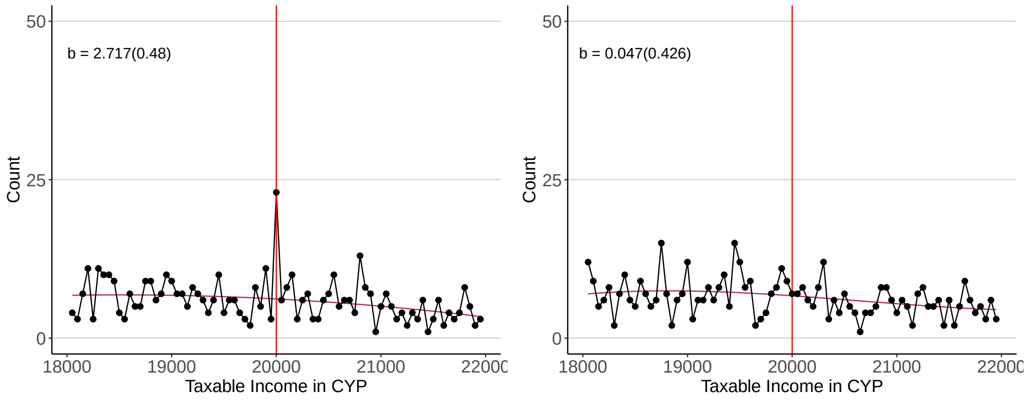
(a) 2003 **(b)** 2004



(c) 2005 **(d)** 2006



(e) 2007 **(f)** 2008



Notes: The figure shows the bunching dynamics of the self-employed around the CYP 20,000 kink between 2003 and 2008. The kink is introduced in 2004, remains in place until 2006 and is removed in 2007. In each panel, the black connected dots plot the empirical distribution in bins of width CYP 50 and the red line is the estimated counterfactual density. Solid vertical lines mark the location of the kink. The panels also show the estimated excess mass b, and for years 2004 to 2006, the observed elasticity e. Bootstrapped standard errors are in parentheses.

Table 1: Summary Statistics

	Mean	Standard Deviation
Gross Income	14,332.29	8,968.91
Deductions	1,630.58	1,610.29
Taxable Income	12,702.72	7,947.56
Tax Bill	891.23	1,743.88
Female	0.395	0.489
Age	40.658	8.100
Wage Earner	0.913	0.282
Sectors		
Agriculture	0.004	0.065
Mining	0.003	0.052
Manufacturing	0.078	0.268
Construction	0.074	0.262
Utilities	0.021	0.144
Trade	0.109	0.311
Services	0.226	0.418
Finance	0.103	0.305
Public	0.369	0.482
Other	0.010	0.098
Observations	1,044,808	

Notes: The table shows summary statistics for our main sample, including individuals aged 25 to 54 in the years 2003 to 2010.

Table 2: Structural Parameter Estimates

	ε	ϕ	ϕ_+	ϕ_-
Panel A: Standard Model, Symmetric Frictions				
CYP 10,000 kink	0.040	54.45		
	(0.003)	(28.49)		
Full set of kinks	0.215	139.52		
	(0.059)	(28.87)		
Panel B: Standard Model, Asymmetric Frictions				
CYP 10,000 kink	0.040		128.22	21.81
	(0.003)		(36.69)	(39.27)
Full set of kinks	0.070		127.30	43.75
	(0.014)		(33.49)	(17.73)
Panel C: Adapted Model, Asymmetric Frictions				
CYP 10,000 kink	0.287		63.21	1.000[†]
	(0.039)		(9.29)	(0.000)
Full set of kinks	0.250		48.69	1.956
	(0.129)		(27.81)	(3.073)

Notes: The table shows structural parameter estimates based on static model specifications. Panel A shows results from the standard model with bunching from above and symmetric frictions, Panel B shows results from the standard model with asymmetric frictions, and Panel C shows results from the adapted model with bunching from below and asymmetric frictions. Each panel presents parameter estimates using the CYP 10,000 kink in the years 2004 to 2008 as empirical bunching moments, as well as estimates using the full set of kinks as bunching moments. The full set includes all kinks from Figures 2 and 3, namely the CYP 10,000 kink in the years 2004 to 2008, the CYP 9,000 kink in the years 2003 to 2005, the CYP 12,000 kink in the years 2003 to 2005, the CYP 10,750 kink in the years 2007 to 2009, and the CYP 11,413 kink in the years 2008 to 2010. [†]We impose a lower bound of 1 on the ϕ parameters in the estimation.

Table 3: Structural Parameter Estimates from Dynamic Models

	ε	ϕ	π_1	$\pi_1\pi_2$	$\pi_1\pi_2\pi_3$
Panel A: Standard Model					
CYP 10,000 kink	0.042	98.08	0.651	0.062	0.002
	(0.155)	(38.80)	(0.070)	(0.144)	(0.063)
Full set of kinks	0.239	143.69	0.623	0.279	0.015
	(0.072)	(11.57)	(0.036)	(0.065)	(0.071)
Panel B: Adapted Model					
CYP 10,000 kink	0.097	11.21	0.075	0.003	0.000
	(0.010)	(1.18)	(0.106)	(0.061)	(0.033)
Full set of kinks	0.132	13.77	0.026	0.001	0.000
	(0.141)	(7.64)	(0.194)	(0.119)	(0.073)

Notes: The table shows structural parameter estimates based on dynamic model specifications. Panel A shows results from the standard model with bunching from above, and Panel B shows results from the adapted model with bunching from below. Both panels present parameter estimates using the CYP 10,000 kink in the years 2004 to 2008 as empirical bunching moments, as well as estimates using the full set of kinks as bunching moments. The full set includes all kinks from Figures 2 and 3, namely the CYP 10,000 kink in the years 2004 to 2008, the CYP 9,000 kink in the years 2003 to 2005, the CYP 12,000 kink in the years 2003 to 2005, the CYP 10,750 kink in the years 2007 to 2009, and the CYP 11,413 kink in the years 2008 to 2010.

Table 4: Structural Parameter Estimates for the Self-Employed

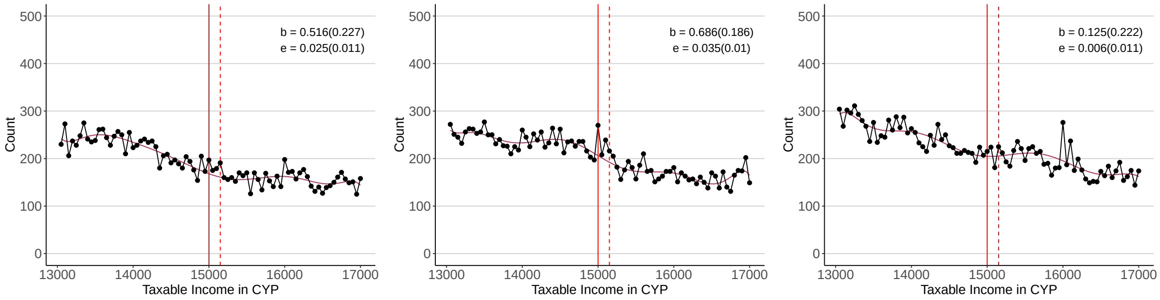
	ε	ϕ	ϕ_+	ϕ_-
Standard model, symmetric frictions	0.297	22.36		
	(0.029)	(5.774)		
Standard model, asymmetric frictions	0.297		6.961	4.956
	(0.029)		(13.87)	(6.729)
Adapted model, asymmetric frictions	0.554		16.05	$1.000^\dagger$
	(0.084)		(3.492)	(0.060)

Notes: The table shows structural parameter estimates for the self-employed, based on static model specifications using the CYP 20,000 kink in the years 2004 to 2008 as empirical bunching moments. Specifications are chosen analogously to Table 2. The title of each row indicates the model used in the estimation and whether frictions are assumed to be symmetric or asymmetric. †We impose a lower bound of 1 on the ϕ parameters in the estimation.

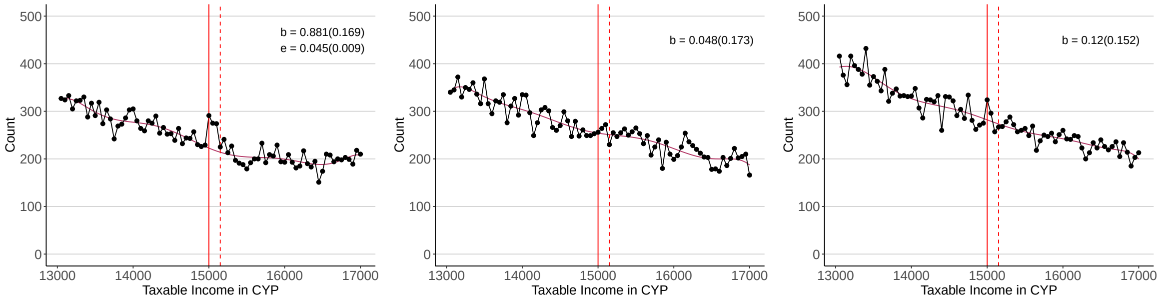
Online Appendix

A Appendix Figures and Tables

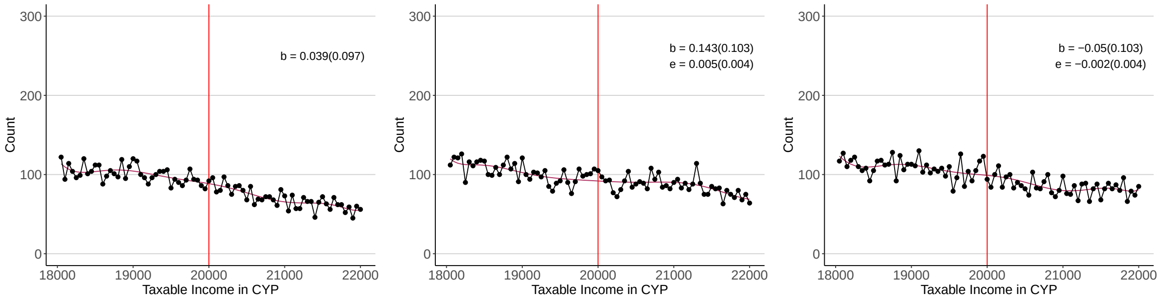
Figure A1: (Lack of) Bunching at Other Tax Kinks



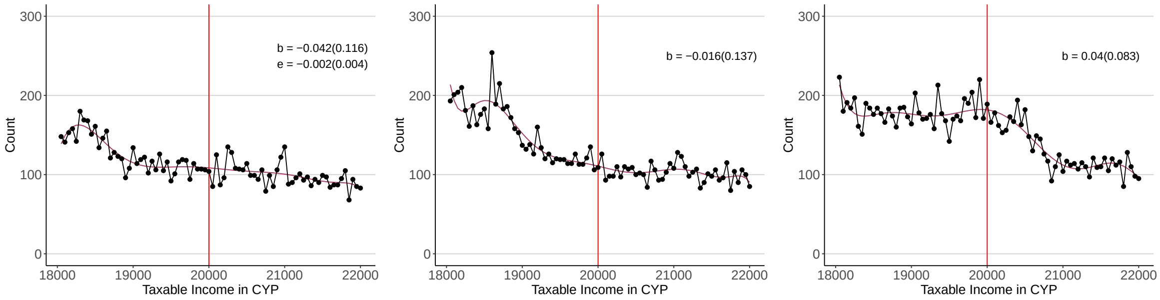
(a) CYP 15,000 kink, 2003 **(b)** CYP 15,000 kink, 2004 **(c)** CYP 15,000 kink, 2005



(d) CYP 15,000 kink, 2006 **(e)** CYP 15,000 kink, 2007 **(f)** CYP 15,000 kink, 2008



(g) CYP 20,000 kink, 2003 **(h)** CYP 20,000 kink, 2004 **(i)** CYP 20,000 kink, 2005

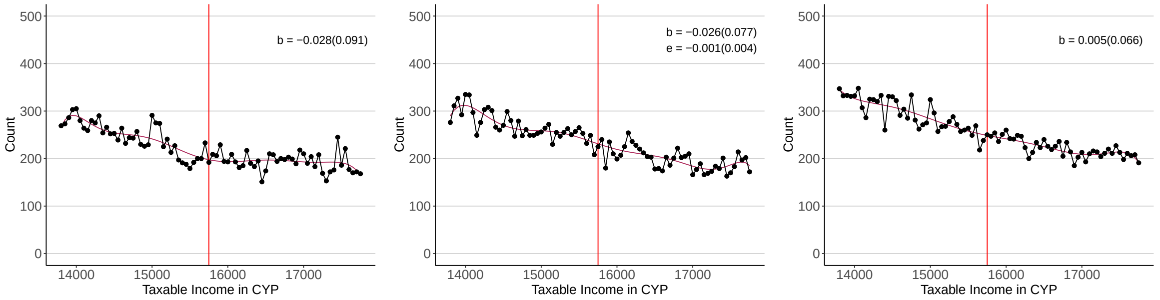


(j) CYP 20,000 kink, 2006 **(k)** CYP 20,000 kink, 2007 **(l)** CYP 20,000 kink, 2008

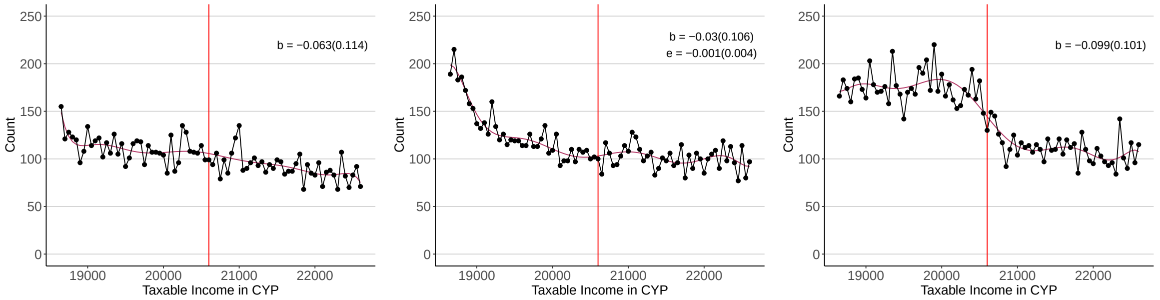
Notes: The figure shows bunching dynamics around different tax kinks. The kink at CYP 15,000 (Panels a to f) is in place since 2003 and removed in 2007. The kink at CYP 20,000 (Panels g to l) is introduced in 2004 and removed in 2007. *(continued on next page)*

Figure A1: (Lack of) Bunching at Other Tax Kinks (Continued)

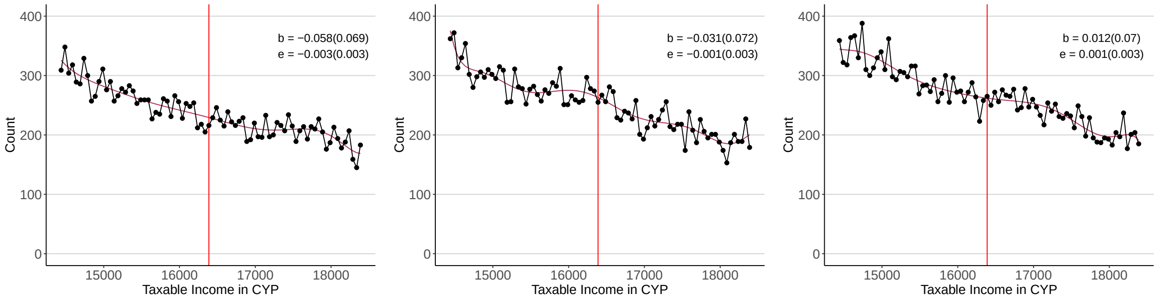
(m) CYP 15,750 kink, 2006 **(n)** CYP 15,750 kink, 2007 **(o)** CYP 15,750 kink, 2008



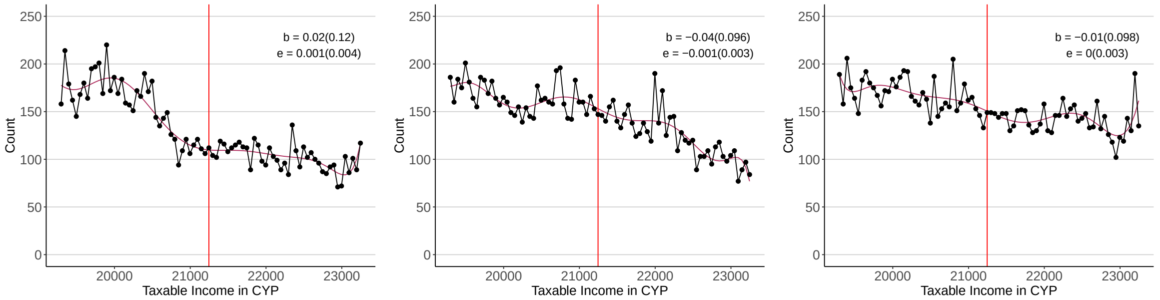
(p) CYP 20,600 kink, 2006 **(q)** CYP 20,600 kink, 2007 **(r)** CYP 20,600 kink, 2008



(s) CYP 16,388 kink, 2008 **(t)** CYP 16,388 kink, 2009 **(u)** CYP 16,388 kink, 2010

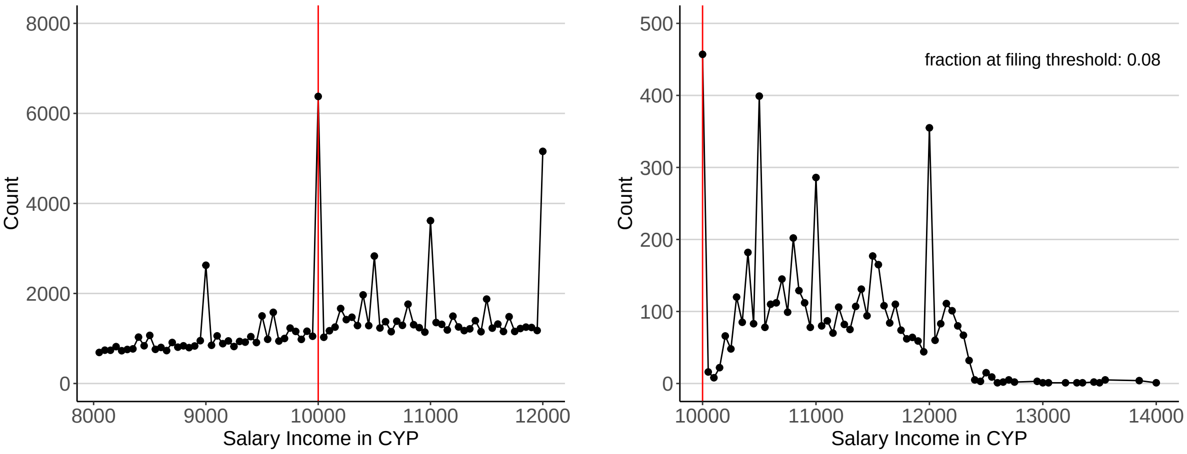


(v) CYP 21,246 kink, 2008 **(w)** CYP 21,246 kink, 2009 **(x)** CYP 21,246 kink, 2010



Notes (continued): The kinks at CYP 15,750 (Panels m to o) and CYP 20,600 (Panels p to r) are introduced in 2007 and removed in 2008. The kinks at CYP 16,388 (Panels s to u) and CYP 21,246 (Panels v to x) are introduced in 2008 and remain in place until 2010. In each panel, the black connected dots plot the empirical distribution in bins of width CYP 50 and the red line is the estimated counterfactual density. Solid vertical lines mark the location of the kink, while dashed lines demarcate the upper bound of the bunching region. The panels also show the estimated excess mass b, and for the years when the respective kink is in place, the observed elasticity e. Bootstrapped standard errors are in parentheses.

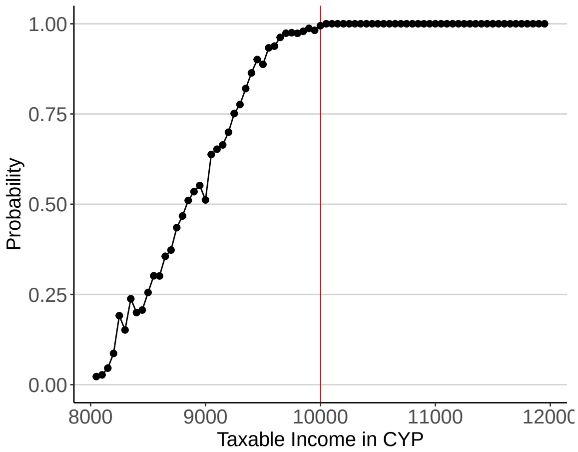
Figure A2: Gross Income Distribution and the Tax Filing Threshold



(a) Gross Income Distribution around the Filing Threshold

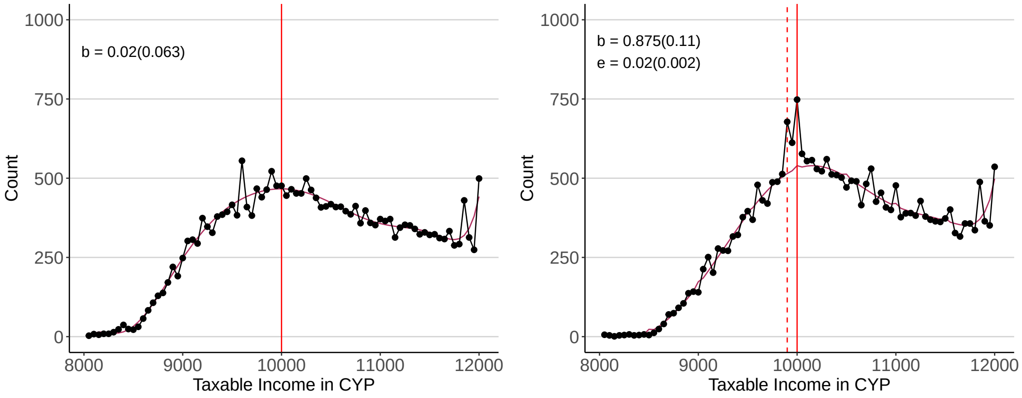
(b) Gross Income Distribution Among Bunchers at the First Tax Kink

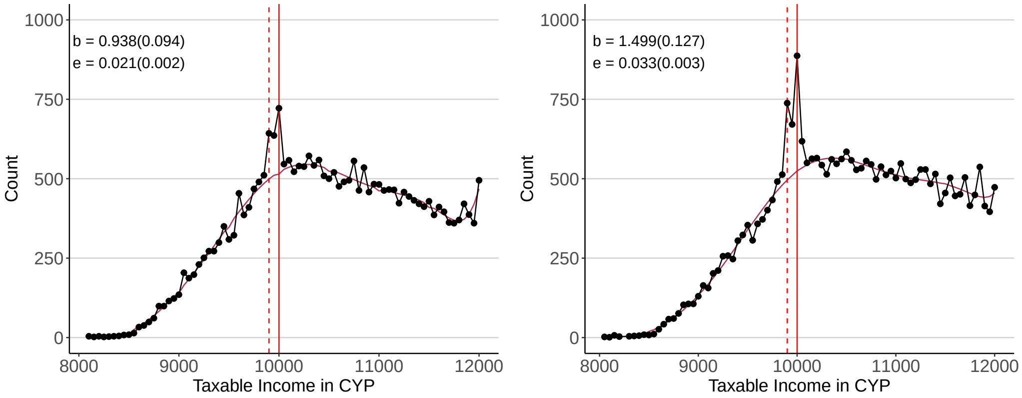
(c) Probability of Being Above the Filing Threshold by Taxable Income

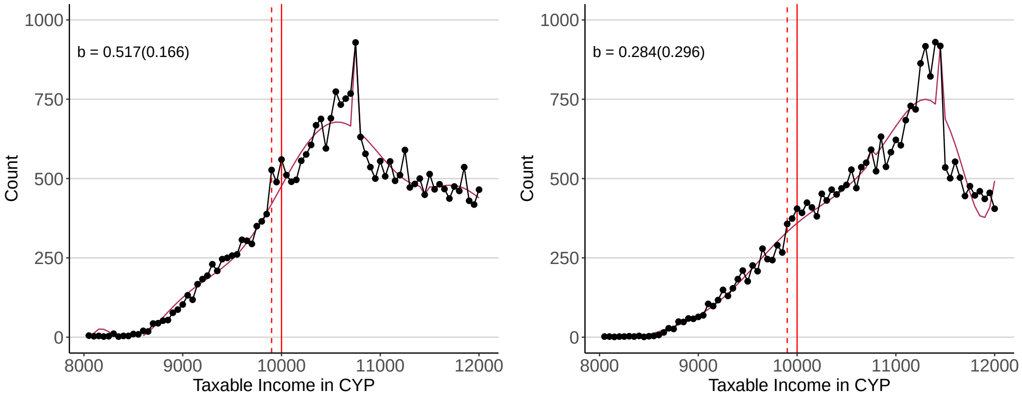


Notes: The figure shows evidence on the role of the tax filing threshold, pooling years 2004 to 2006, when the first tax kink is at CYP 10,000 taxable income and the filing threshold is at CYP 10,000 gross income. Panel (a) shows the distribution of gross income around the filing threshold, above which taxpayers are mandated to file a tax return. Panel (b) shows the distribution of gross income among individuals who bunch in terms of taxable income at the first tax kink. “Fraction at filing threshold” denotes the fraction of taxable income bunchers whose gross income is exactly at the filing threshold. Panel (c) shows the probability of being above the filing threshold, that is the probability of having gross income of at least CYP 10,000, by taxable income. In all panels, bin width is CYP 50, and the vertical red line marks the location of the filing threshold.

Figure A3: Bunching at the CYP 10,000 Kink, Excluding Individuals at the Filing Threshold


(a) 2003 **(b)** 2004

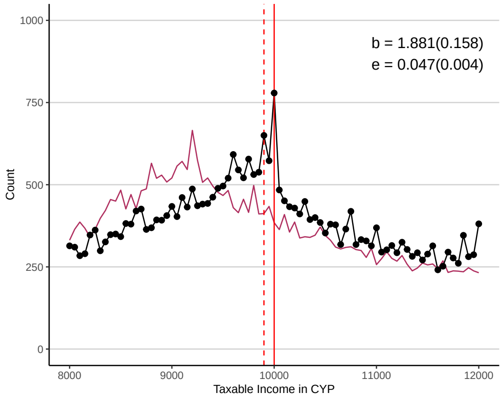

(c) 2005 **(d)** 2006


(e) 2007 **(f)** 2008

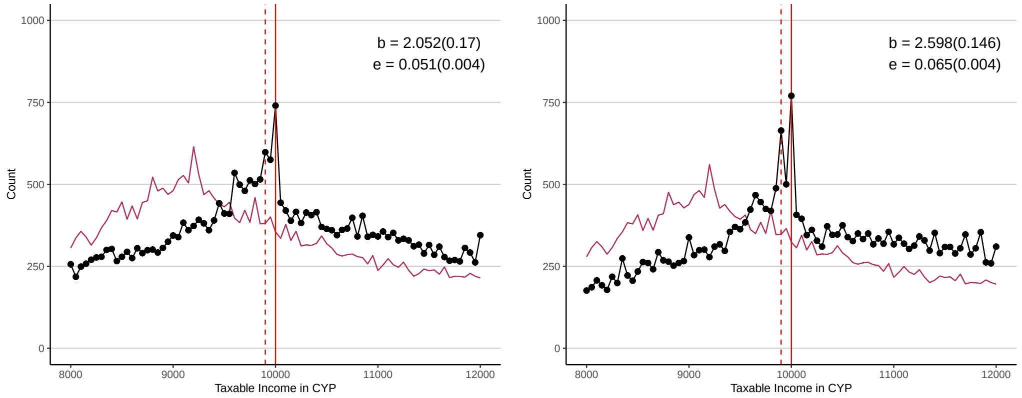
Notes: The figure shows bunching at the CYP 10,000 kink between 2003 and 2008 analogous to Figure 2, but excluding individuals with gross income at the tax filing threshold. In each panel, the black connected dots plot the empirical distribution in bins of width CYP 50 and the red line is the estimated counterfactual density. Solid vertical lines mark the location of the kink, while dashed lines demarcate the lower bound of the bunching region. The panels also show the estimated excess mass *b*, and for years 2004 to 2006, the observed elasticity *e*. Bootstrapped standard errors are in parentheses.

Figure A4: Bunching at the CYP 10,000 Kink, Using the 2003 Density as Counterfactual

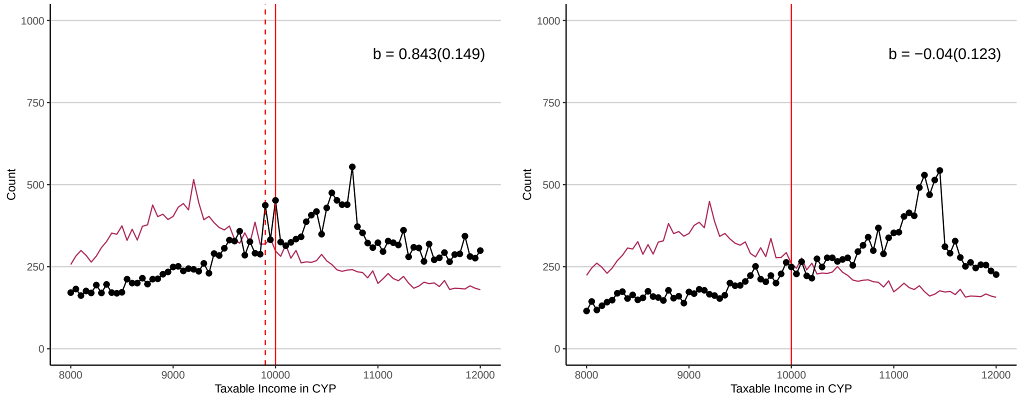
(a) 2004



(b) 2005 **(c)** 2006



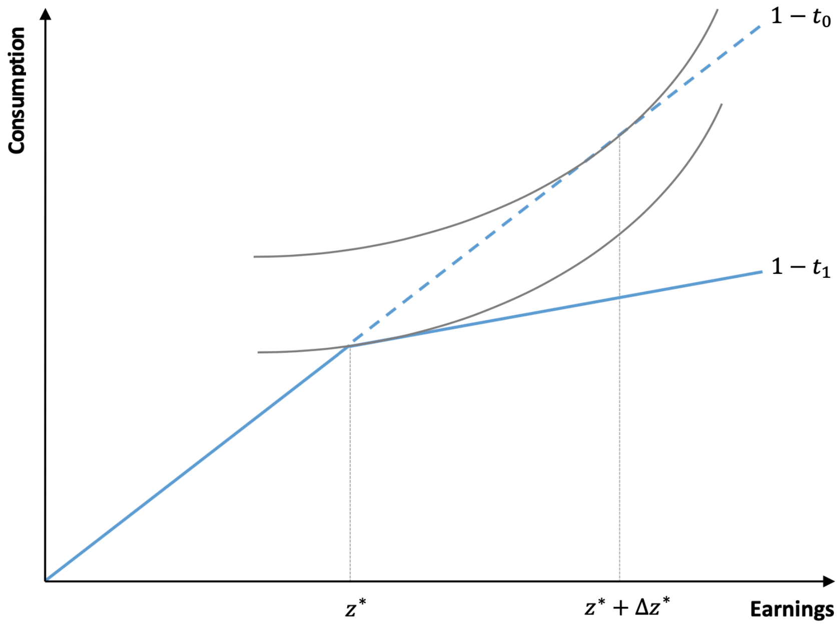
(d) 2007 **(e)** 2008



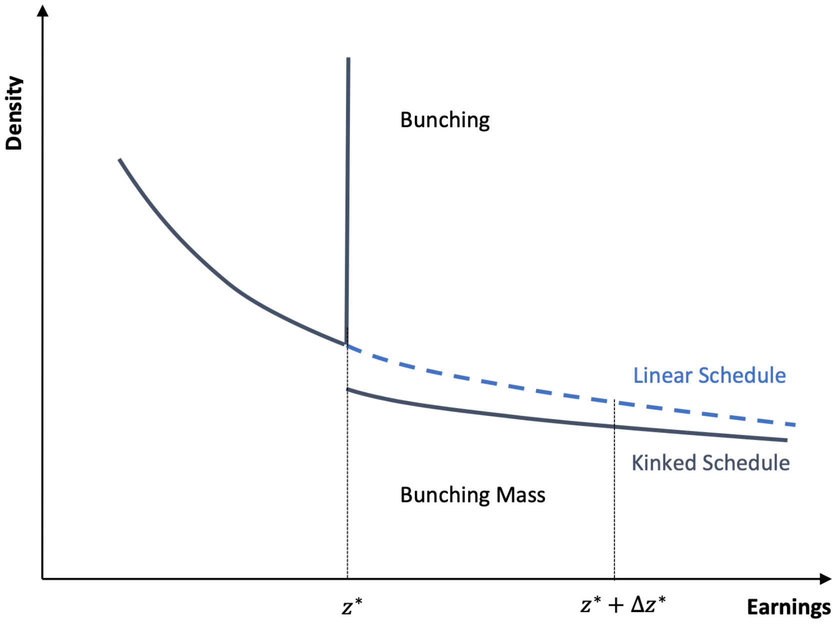
Notes: The figure shows bunching at the CYP 10,000 kink between 2004 and 2008 analogous to Figure 2, but using an alternative counterfactual density. The counterfactual is based on the empirical density observed in 2003, adjusted for inflation and normalized to match the sample size within the bandwidth in each year. In each panel, the black connected dots plot the empirical distribution in bins of width CYP 50 by year and the red line is the counterfactual density. Solid vertical lines mark the location of the kink, while dashed lines demarcate the lower bound of the bunching region. The panels also show the estimated excess mass b, and for years 2004 to 2006, the observed elasticity e. Bootstrapped standard errors are in parentheses.

Figure A5: Bunching at a Kink in a Frictionless World

(a) Optimization with a Kink

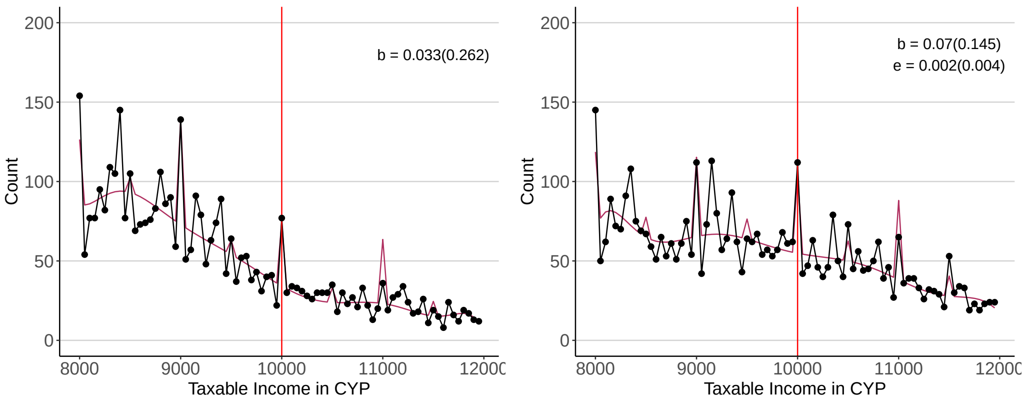


(b) Taxable Income Distribution

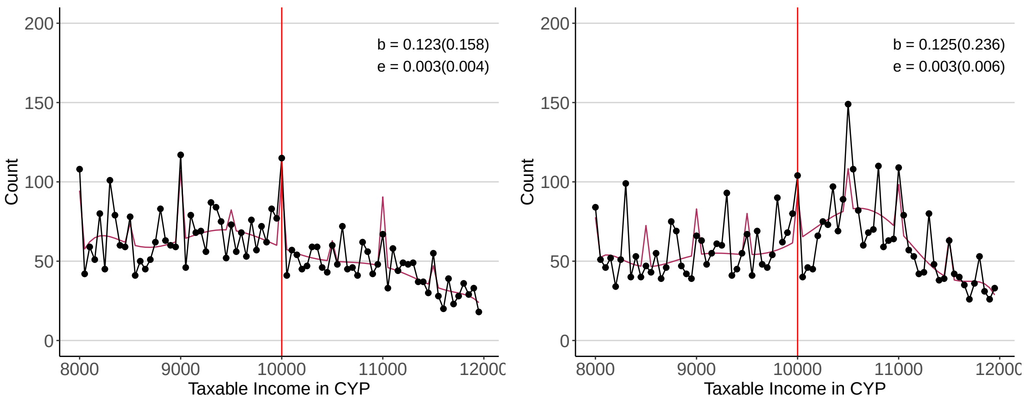


Notes: The figure depicts the predicted effect of introducing a convex kink at earnings level z^* on the distribution of taxable income under the model without frictions from Section 4.1. Panel (a) presents budget sets under a linear tax schedule and under a kinked schedule, as well as indifference curves of the marginal bunching individual. Panel (b) shows the predicted distribution of taxable income, featuring bunching at the kink.

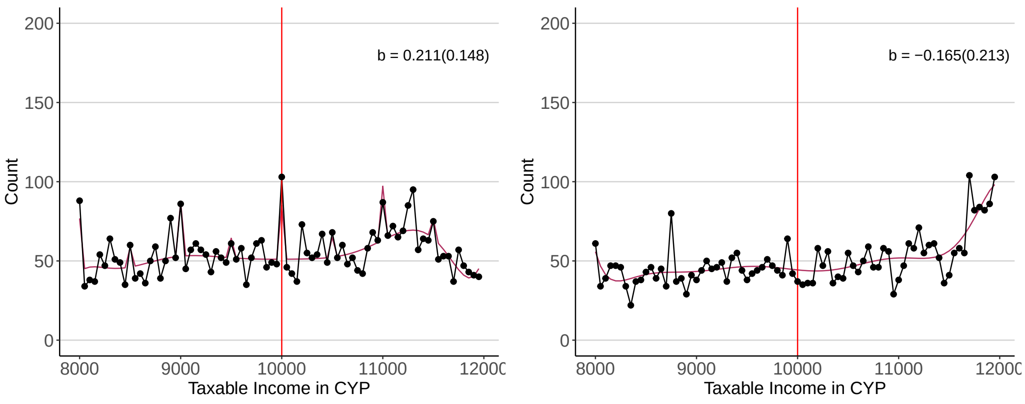
Figure A6: Bunching of the Self-Employed at the CYP 10,000 Kink



(a) 2003 **(b)** 2004



(c) 2005 **(d)** 2006



(e) 2007 **(f)** 2008

Notes: The figure shows the bunching dynamics of the self-employed around the CYP 10,000 kink between 2003 and 2008. The kink is introduced in 2004, remains in place until 2006 and is removed in 2007. In each panel, the black connected dots plot the empirical distribution in bins of width CYP 50 and the red line is the estimated counterfactual density. Solid vertical lines mark the location of the kink. The panels also show the estimated excess mass b, and for years 2004 to 2006, the observed elasticity e. Bootstrapped standard errors are in parentheses.

Table A1: Bunching Estimates and Robustness at the CYP 10,000 Kink

	Excess Mass						Observed Elasticity		
	b_{2003}	b_{2004}	b_{2005}	b_{2006}	b_{2007}	b_{2008}	ε_{2004}	ε_{2005}	ε_{2006}
Main estimates	0.060	1.036	1.158	1.777	0.486	0.095	0.026	0.029	0.044
	(0.069)	(0.113)	(0.103)	(0.121)	(0.156)	(0.244)	(0.003)	(0.003)	(0.003)
Robustness checks									
Bandwidth: 1000	0.069	1.174	1.220	1.615	0.735	0.084	0.029	0.031	0.040
	(0.072)	(0.125)	(0.134)	(0.112)	(0.130)	(0.066)	(0.003)	(0.003)	(0.003)
Bin Size: 100	0.022	0.600	0.738	1.062	0.232	0.043	0.030	0.037	0.053
	(0.042)	(0.089)	(0.077)	(0.102)	(0.106)	(0.104)	(0.004)	(0.004)	(0.005)
Polynomial Order: 6	0.072	1.169	1.291	1.950	0.351	0.138	0.029	0.032	0.049
	(0.074)	(0.141)	(0.143)	(0.148)	(0.157)	(0.168)	(0.004)	(0.004)	(0.004)
Excluded Region -1 below	0.035	0.650	0.829	1.115	0.224	0.098	0.016	0.021	0.028
	(0.104)	(0.108)	(0.098)	(0.133)	(0.125)	(0.190)	(0.003)	(0.002)	(0.003)
Excluded Region -3 below	0.144	1.051	1.258	1.999	0.459	-0.023	0.026	0.031	0.050
	(0.154)	(0.138)	(0.124)	(0.142)	(0.193)	(0.289)	(0.003)	(0.003)	(0.004)
Excluded Region -4 below	0.106	1.004	1.260	2.074	0.396	-0.052	0.025	0.031	0.052
	(0.183)	(0.157)	(0.143)	(0.170)	(0.220)	(0.334)	(0.004)	(0.004)	(0.004)
Counterfactual Shifted Above Kink	0.049	0.972	1.089	1.665	0.419	0.061	0.024	0.027	0.042
	(0.077)	(0.214)	(0.227)	(0.333)	(0.165)	(0.123)	(0.005)	(0.006)	(0.008)
Counterfactual Shifted Below Kink	0.049	0.925	1.026	1.544	0.396	0.060	0.023	0.026	0.039
	(0.077)	(0.202)	(0.209)	(0.297)	(0.156)	(0.121)	(0.005)	(0.005)	(0.007)
2003 Density as Counterfactual		1.881	2.052	2.598	0.843	-0.040	0.047	0.051	0.065
		(0.158)	(0.170)	(0.146)	(0.149)	(0.123)	(0.004)	(0.004)	(0.004)
Observations	41,310	41,873	41,307	42,933	39,896	35,678	41,873	41,307	42,933

Notes: The table shows estimates of the yearly excess mass and, for those years the kink is in place, of the observed elasticity at the CYP 10,000 kink. Besides our main estimates, the table shows robustness to a different bandwidth, bin size, polynomial order and to different choices of the excluded region (z_L) from those used in the main specification. Moreover, the last three rows show results from alternative counterfactual estimation procedures. The third-to-last row shows estimates based on a polynomial counterfactual density as the main specification, but with an additional shift to account for bunching originating from above. The second-to-last row shows estimates under an analogously shifted counterfactual accounting for bunching from below. The last row shows estimates based on an alternative counterfactual density based on the empirically observed distribution in 2003, the year before the kink is introduced. See Appendix Figure A4 for graphs and details of the previous-year counterfactual specification. "Observations" denotes number of individuals in the main estimation bandwidth. Bootstrapped standard errors are shown in parentheses.

Table A2: Summary Statistics for Wage Earners Around the First Kink vs. Above

	(1) All Wage Earners		(2) Around First Kink		(3) Above First Kink	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Gross Income	14,580.02	8,947.56	11,805.41	1,757.86	18,525.55	3,472.70
Deductions	1,662.30	1,562.97	1,388.79	836.47	2,160.42	1,281.12
Taxable Income	12,918.79	7,931.13	10,416.62	1,353.31	16,365.12	2,949.58
Tax Bill	924.47	1,735.19	84.45	121.20	1,240.62	658.05
Female	0.405	0.491	0.325	0.469	0.391	0.488
Age	40.383	8.113	39.006	8.172	39.966	7.678
Sectors						
Agriculture	0.004	0.065	0.007	0.081	0.002	0.044
Mining	0.003	0.054	0.003	0.056	0.003	0.057
Manufacturing	0.078	0.268	0.114	0.318	0.052	0.222
Construction	0.074	0.262	0.112	0.316	0.049	0.217
Utilities	0.021	0.144	0.021	0.145	0.024	0.153
Trade	0.109	0.311	0.148	0.355	0.061	0.238
Services	0.226	0.418	0.251	0.434	0.175	0.380
Finance	0.104	0.305	0.073	0.261	0.121	0.326
Public	0.369	0.483	0.258	0.438	0.497	0.500
Other	0.010	0.098	0.009	0.095	0.014	0.117
Observations	953,586		326,916		298,944	

Notes: The table shows summary statistics for wage earners around the first tax kink vs. at higher income levels. Column (1) shows summary statistics for all wage earners. Column (2) focuses on wage earners within the bandwidth around the first tax kink, i.e. within +/- CYP 2000 of the location of the first kink in a given year. Column (3) includes wage earners above the bandwidth of the first kink. All columns apply the same general sample restrictions as Table 1, including individuals aged 25 to 54 in the years 2003 to 2010.

Table A3: Extended Structural Parameter Estimates

	ε	ϕ	ϕ_+	ϕ_-
Panel A: Standard Model, Symmetric Frictions				
CYP 10,000 kink	0.040	54.45		
	(0.003)	(28.49)		
Intermediate set of kinks	0.069	64.48		
	(0.009)	(40.23)		
Full set of kinks	0.215	139.52		
	(0.059)	(28.87)		
Panel B: Standard Model, Asymmetric Frictions				
CYP 10,000 kink	0.040		128.22	21.81
	(0.003)		(36.69)	(39.27)
Intermediate set of kinks	0.069		196.29	152.44
	(0.009)		(45.13)	(40.00)
Full set of kinks	0.070		127.30	43.75
	(0.014)		(33.49)	(17.73)
Panel C: Adapted Model, Symmetric Frictions				
CYP 10,000 kink	0.040	$1.000^\dagger$		
	(0.003)	(0.003)		
Intermediate set of kinks	0.069	7.443		
	(0.009)	(1.984)		
Full set of kinks	0.069	$1.000^\dagger$		
	(0.009)	(0.031)		
Panel D: Adapted Model, Asymmetric Frictions				
CYP 10,000 kink	0.287		63.21	$1.000^\dagger$
	(0.039)		(9.29)	(0.000)
Intermediate set of kinks	0.341		74.68	1.046
	(0.116)		(27.82)	(2.686)
Full set of kinks	0.250		48.69	1.956
	(0.129)		(27.81)	(3.073)

Notes: The table shows extended structural parameter estimates based on static model specifications. Panel A shows results from the standard model with bunching from above and symmetric frictions, Panel B shows results from the standard model with asymmetric frictions, Panel C shows results from the adapted model with bunching from below and symmetric frictions, and Panel D shows results from the adapted model with asymmetric frictions. Each panel presents parameter estimates using the CYP 10,000 kink in the years 2004 to 2008 as empirical bunching moments, as well as estimates using an intermediate set of kinks and the full set of kinks as bunching moments. The full set includes all kinks from Figures 2 and 3, namely the CYP 10,000 kink in the years 2004 to 2008, the CYP 9,000 kink in the years 2003 to 2005, the CYP 12,000 kink in the years 2003 to 2005, the CYP 10,750 kink in the years 2007 to 2009, and the CYP 11,413 kink in the years 2008 to 2010. The intermediate set of kinks corresponds to the full set excluding the CYP 9,000 and CYP 10,750 kinks. †We impose a lower bound of 1 on the ϕ parameters in the estimation.

Table A4: Extended Structural Parameter Estimates from Dynamic Models

	ε	ϕ	ϕ_+	ϕ_-	π_1	$\pi_1\pi_2$	$\pi_1\pi_2\pi_3$
Panel A: Standard Model, Symmetric Frictions							
CYP 10,000 kink	0.042	98.08			0.651	0.062	0.002
	(0.155)	(38.80)			(0.070)	(0.144)	(0.063)
Intermediate set of kinks	0.091	133.36			0.667	0.444	0.236
	(0.141)	(37.28)			(0.076)	(0.167)	(0.074)
Full set of kinks	0.239	143.69			0.623	0.279	0.015
	(0.072)	(11.57)			(0.036)	(0.065)	(0.071)
Panel B: Standard Model, Asymmetric Frictions							
CYP 10,000 kink	0.041		112.01	39.73	0.654	0.401	0.024
	(0.120)		(38.02)	(42.66)	(0.037)	(0.072)	(0.067)
Intermediate set of kinks	0.069		52.72	12.00	0.348	0.121	0.002
	(0.131)		(50.24)	(59.40)	(0.129)	(0.139)	(0.072)
Full set of kinks	0.138		129.06	17.38	0.518	0.161	0.008
	(0.131)		(34.19)	(43.10)	(0.036)	(0.070)	(0.065)
Panel C: Adapted Model, Symmetric Frictions							
CYP 10,000 kink	0.097	11.21			0.075	0.003	0.000
	(0.010)	(1.18)			(0.106)	(0.061)	(0.033)
Intermediate set of kinks	0.127	12.81			0.149	0.022	0.003
	(0.152)	(9.50)			(0.214)	(0.145)	(0.094)
Full set of kinks	0.132	13.77			0.026	0.001	0.000
	(0.141)	(7.64)			(0.194)	(0.119)	(0.073)
Panel D: Adapted Model, Asymmetric Frictions							
CYP 10,000 kink	0.051		9.857	1.830	0.471	0.222	0.004
	(0.164)		(25.733)	(11.687)	(0.170)	(0.149)	(0.089)
Intermediate set of kinks	0.448		14.72	9.744	0.047	0.002	0.000
	(0.152)		(14.67)	(11.058)	(0.120)	(0.070)	(0.041)
Full set of kinks	0.450		31.10	31.09	0.010	0.000	0.000
	(0.129)		(12.63)	(14.65)	(0.041)	(0.007)	(0.002)

Notes: The table shows extended structural parameter estimates based on dynamic model specifications. Panel A shows results from the standard model with bunching from above and symmetric frictions, Panel B shows results from the standard model with asymmetric frictions, Panel C shows results from the adapted model with bunching from below and symmetric frictions, and Panel D shows results from the adapted model with asymmetric frictions. Each panel presents parameter estimates using the CYP 10,000 kink in the years 2004 to 2008 as empirical bunching moments, as well as estimates using an intermediate set of kinks and the full set of kinks as bunching moments. The full set includes all kinks from Figures 2 and 3, namely the CYP 10,000 kink in the years 2004 to 2008, the CYP 9,000 kink in the years 2003 to 2005, the CYP 12,000 kink in the years 2003 to 2005, the CYP 10,750 kink in the years 2007 to 2009, and the CYP 11,413 kink in the years 2008 to 2010. The intermediate set of kinks corresponds to the full set excluding the CYP 9,000 and CYP 10,750 kinks.

Table A5: Goodness of Fit

	(1) CYP 10,000 kink	(2) Full set of kinks
Static Models (Results from Table 2)		
Standard Model, Symmetric Frictions	0.02184	1.02205
Standard Model, Asymmetric Frictions	0.02183	0.04736
Adapted Model, Asymmetric Frictions	0.06618	0.19303
Dynamic Models (Results from Table 3)		
Standard Model	0.07528	1.70606
Adapted Model	0.00895	0.00990

Notes: The table presents a comparison of fit of the models we estimate in Tables 2 and 3. Each row corresponds to a model specification and each column corresponds to a set of bunching moments used in the estimation. For each model and set of moments, the table shows the sum of squared estimation errors, i.e. the objective function from the structural estimation, normalized by the sum of squared data moments. All values are expressed in percent.

B Full Bunching and Robustness Results

Table B1: Bunching Estimates and Robustness at the CYP 9,000 Kink

	b_{2003}	b_{2004}	b_{2005}	ε_{2003}
Main estimates	0.314	0.062	0.040	0.009
	(0.054)	(0.072)	(0.085)	(0.002)
Robustness checks				
Bandwidth: 1000	0.307	0.086	0.040	0.009
	(0.057)	(0.074)	(0.068)	(0.002)
Bin Size: 100	0.138	0.051	0.027	0.008
	(0.034)	(0.058)	(0.064)	(0.002)
Polynomial Order: 6	0.329	0.059	0.022	0.009
	(0.060)	(0.082)	(0.103)	(0.002)
Excluded Region -1 below	0.287	0.089	0.029	0.008
	(0.077)	(0.109)	(0.129)	(0.002)
Excluded Region -2 below	0.425	0.124	0.085	0.012
	(0.098)	(0.144)	(0.168)	(0.003)
Counterfactual Shifted Above Kink	0.307	0.061	0.040	0.009
	(0.072)	(0.072)	(0.084)	(0.002)
Counterfactual Shifted Below Kink	0.307	0.061	0.039	0.009
	(0.071)	(0.072)	(0.083)	(0.002)
Observations	43,993	41,827	39,574	43,993

Notes: The table shows estimates of the yearly excess mass and, for those years the kink is in place, of the observed elasticity at the CYP 9,000 kink. Besides our main estimates, the table shows robustness to a different bandwidth, bin size, polynomial order and to different choices of the excluded region (z_L) from those used in the main specification. Moreover, the last two rows show results from alternative counterfactual estimation procedures. The second-to-last row shows estimates based on a polynomial counterfactual density as the main specification, but with an additional shift to account for bunching originating from above. The last row shows estimates under an analogously shifted counterfactual accounting for bunching from below. "Observations" denotes number of individuals in the main estimation bandwidth. Bootstrapped standard errors are shown in parentheses.

Table B2: Bunching Estimates and Robustness at the CYP 12,000 Kink

	b_{2003}	b_{2004}	b_{2005}	ε_{2003}
Main estimates	1.041	1.070	0.646	0.069
	(0.136)	(0.143)	(0.141)	(0.009)
Robustness checks				
Bandwidth: 1000	1.015	1.116	0.684	0.065
	(0.161)	(0.192)	(0.161)	(0.010)
Bin Size: 100	0.525	0.570	0.328	0.068
	(0.066)	(0.058)	(0.062)	(0.008)
Polynomial Order: 6	0.953	1.105	0.578	0.061
	(0.124)	(0.138)	(0.134)	(0.008)
Excluded Region -1 below	0.484	0.461	0.366	0.031
	(0.106)	(0.109)	(0.091)	(0.007)
Excluded Region -2 below	0.480	0.497	0.437	0.031
	(0.138)	(0.141)	(0.116)	(0.009)
Excluded Region -4 below	0.965	1.038	0.645	0.062
	(0.153)	(0.172)	(0.170)	(0.010)
Counterfactual Shifted Above Kink	0.973	0.996	0.607	0.063
	(0.262)	(0.227)	(0.181)	(0.017)
Counterfactual Shifted Below Kink	0.989	1.016	0.616	0.066
	(0.267)	(0.230)	(0.181)	(0.018)
Observations	25,559	28,703	30,969	25,559

Notes: The table shows estimates of the yearly excess mass and, for those years the kink is in place, of the observed elasticity at the CYP 12,000 kink. Besides our main estimates, the table shows robustness to a different bandwidth, bin size, polynomial order and to different choices of the excluded region (z_L) from those used in the main specification. Moreover, the last two rows show results from alternative counterfactual estimation procedures. The second-to-last row shows estimates based on a polynomial counterfactual density as the main specification, but with an additional shift to account for bunching originating from above. The last row shows estimates under an analogously shifted counterfactual accounting for bunching from below. "Observations" denotes number of individuals in the main estimation bandwidth. Bootstrapped standard errors are shown in parentheses.

Table B3: Bunching Estimates and Robustness at the CYP 10,750 Kink

	b_{2006}	b_{2007}	b_{2008}	ε_{2007}
Main estimates	0.080	1.459	0.043	0.034
	(0.116)	(0.192)	(0.347)	(0.004)
Robustness checks				
Bandwidth: 1000	0.035	1.006	0.279	0.023
	(0.072)	(0.191)	(0.368)	(0.004)
Bin Size: 100	0.104	1.137	-0.161	0.053
	(0.138)	(0.188)	(0.337)	(0.009)
Polynomial Order: 6	0.044	1.589	-0.440	0.037
	(0.120)	(0.186)	(0.349)	(0.004)
Excluded Region -1 below	0.183	0.674	-0.005	0.016
	(0.175)	(0.115)	(0.185)	(0.003)
Excluded Region -2 below	0.221	0.898	-0.024	0.021
	(0.232)	(0.140)	(0.224)	(0.003)
Excluded Region -3 below	0.221	1.114	-0.165	0.026
	(0.277)	(0.165)	(0.279)	(0.004)
Counterfactual Shifted Above Kink	0.079	1.331	0.040	0.031
	(0.115)	(0.294)	(0.325)	(0.007)
Counterfactual Shifted Below Kink	0.077	1.137	0.035	0.026
	(0.113)	(0.236)	(0.275)	(0.005)
Observations	42,530	41,003	37,739	41,003

Notes: The table shows estimates of the yearly excess mass and, for those years the kink is in place, of the observed elasticity at the CYP 10,750 kink. Besides our main estimates, the table shows robustness to a different bandwidth, bin size, polynomial order and to different choices of the excluded region (z_L) from those used in the main specification. Moreover, the last two rows show results from alternative counterfactual estimation procedures. The second-to-last row shows estimates based on a polynomial counterfactual density as the main specification, but with an additional shift to account for bunching originating from above. The last row shows estimates under an analogously shifted counterfactual accounting for bunching from below. "Observations" denotes number of individuals in the main estimation bandwidth. Bootstrapped standard errors are shown in parentheses.

Table B4: Bunching Estimates and Robustness at the CYP 11,413 Kink

	b_{2007}	b_{2008}	b_{2009}	b_{2010}	ε_{2008}	ε_{2009}	ε_{2010}
Main estimates	0.034	2.313	1.938	2.104	0.051	0.042	0.046
	(0.108)	(0.163)	(0.139)	(0.151)	(0.004)	(0.003)	(0.003)
Robustness checks							
Bandwidth: 1000	-0.029	1.472	1.100	1.446	0.032	0.024	0.032
	(0.086)	(0.158)	(0.141)	(0.167)	(0.003)	(0.003)	(0.004)
Bin Size: 100	0.020	1.616	1.408	1.737	0.071	0.062	0.076
	(0.082)	(0.137)	(0.122)	(0.118)	(0.006)	(0.005)	(0.005)
Polynomial Order: 6	-0.017	2.732	2.404	2.389	0.060	0.053	0.052
	(0.109)	(0.206)	(0.195)	(0.187)	(0.005)	(0.004)	(0.004)
Excluded Region -1 below	0.037	1.049	0.918	1.090	0.023	0.020	0.024
	(0.150)	(0.128)	(0.104)	(0.109)	(0.003)	(0.002)	(0.002)
Excluded Region -2 below	-0.038	1.437	1.201	1.432	0.031	0.026	0.031
	(0.190)	(0.147)	(0.119)	(0.123)	(0.003)	(0.003)	(0.003)
Excluded Region -3 below	0.178	2.159	1.709	1.879	0.047	0.037	0.041
	(0.230)	(0.138)	(0.125)	(0.139)	(0.003)	(0.003)	(0.003)
Counterfactual Shifted Above Kink	0.034	2.084	1.753	1.906	0.046	0.038	0.042
	(0.108)	(0.437)	(0.373)	(0.406)	(0.010)	(0.008)	(0.009)
Counterfactual Shifted Below Kink	0.033	1.747	1.476	1.593	0.038	0.032	0.035
	(0.105)	(0.333)	(0.286)	(0.308)	(0.007)	(0.006)	(0.007)
Observations	40,002	38,876	37,742	38,917	38,876	37,742	38,917

Notes: The table shows estimates of the yearly excess mass and, for those years the kink is in place, of the observed elasticity at the CYP 11,413 kink. Besides our main estimates, the table shows robustness to a different bandwidth, bin size, polynomial order and to different choices of the excluded region (z_L) from those used in the main specification. Moreover, the last two rows show results from alternative counterfactual estimation procedures. The second-to-last row shows estimates based on a polynomial counterfactual density as the main specification, but with an additional shift to account for bunching originating from above. The last row shows estimates under an analogously shifted counterfactual accounting for bunching from below. "Observations" denotes number of individuals in the main estimation bandwidth. Bootstrapped standard errors are shown in parentheses.

Table B5: Bunching Estimates and Robustness among Self-Employed at the CYP 20,000 Kink

	b_{2003}	b_{2004}	b_{2005}	b_{2006}	b_{2007}	b_{2008}	ε_{2004}	ε_{2005}	ε_{2006}
Main estimates	1.023	3.506	6.486	8.272	2.717	0.047	0.131	0.243	0.310
	(0.581)	(0.672)	(0.774)	(0.792)	(0.480)	(0.426)	(0.025)	(0.029)	(0.030)
Robustness checks									
Bandwidth: 1000	1.024	3.645	8.698	8.893	2.958	-0.019	0.137	0.326	0.333
	(0.696)	(0.817)	(1.596)	(1.026)	(0.633)	(0.372)	(0.031)	(0.060)	(0.038)
Bin Size: 100	0.765	2.274	3.215	4.332	1.162	0.294	0.171	0.241	0.325
	(0.450)	(0.612)	(0.569)	(0.494)	(0.332)	(0.292)	(0.046)	(0.043)	(0.037)
Polynomial Order: 6	0.901	4.052	8.603	9.115	2.797	0.009	0.152	0.323	0.342
	(0.571)	(1.014)	(1.755)	(1.221)	(0.561)	(0.421)	(0.038)	(0.066)	(0.046)
Excluded Region -1 below	0.379	4.178	6.999	9.385	2.123	0.408	0.157	0.262	0.352
	(0.792)	(0.981)	(1.034)	(1.057)	(0.609)	(0.585)	(0.037)	(0.039)	(0.040)
Excluded Region -2 below	1.041	4.073	6.898	8.787	2.985	1.130	0.153	0.259	0.330
	(1.059)	(1.119)	(1.244)	(1.131)	(0.775)	(0.748)	(0.042)	(0.047)	(0.042)
Counterfactual Shifted Above Kink	0.978	3.300	5.889	7.271	2.565	0.047	0.124	0.221	0.273
	(0.607)	(0.845)	(1.530)	(2.365)	(0.554)	(0.389)	(0.032)	(0.057)	(0.089)
Counterfactual Shifted Below Kink	0.999	3.315	5.905	7.520	2.599	0.047	0.124	0.221	0.282
	(0.587)	(0.813)	(1.527)	(2.413)	(0.550)	(0.419)	(0.030)	(0.057)	(0.090)
Observations	160	252	346	434	476	503	252	346	434

Notes: The table shows estimates of the yearly excess mass and, for those years the kink is in place, of the observed elasticity among the self-employed at the CYP 20,000 kink. Besides our main estimates, the table shows robustness to a different bandwidth, bin size, polynomial order and to different choices of the excluded region (z_L) from those used in the main specification. Moreover, the last two rows show results from alternative counterfactual estimation procedures. The second-to-last row shows estimates based on a polynomial counterfactual density as the main specification, but with an additional shift to account for bunching originating from above. The last row shows estimates under an analogously shifted counterfactual accounting for bunching from below. "Observations" denotes number of individuals in the main estimation bandwidth. Bootstrapped standard errors are shown in parentheses.

C Empirical Methodology

C.1 Reduced-Form Bunching Estimation

We estimate bunching[23] following Chetty et al. (2011) where a counterfactual is fitted to the observed distribution of taxable income around each kink, excluding the bins in the bunching region around the kink itself. The counterfactual c_j is estimated by grouping individuals into earnings bins j of width δ and fitting a flexible polynomial to the empirical distribution. Bunching can then be estimated via the following regression:

$$c_j = \sum_{i=0}^{p} \beta_i (z_j)^i + \sum_{i=z_L}^{z_U} \gamma_i \mathbb{1}[z_j = i] + \sum_{r \in R} \rho_r \mathbb{1}\Big[\frac{z_j}{r} \in \mathbb{N}\Big] + \sum_{k \in K} \theta_k \mathbb{1}\Big[z_j \in K \wedge z_j \notin [z_L, z_U]\Big] + v_j \quad (14)$$

where c_j is the count of individuals in bin j, z_j is the earnings level in the bin, and p is the order of the fitted polynomial. Our main specification uses $\delta = 50$ and $p = 9$. To account for potential diffuse bunching, we define the bunching region to be $z \in [z_L, z_U]$ around a kink at z^*, where z_L and z_U are determined visually at each kink. If there is no visible diffuse bunching below (above) a kink, z_L (z_U) is set to z^*, the location of the kink itself. Because kinks are located at round numbers, and those may act as reference points, we also control for round-number bunching. This avoids over-estimating bunching responses at a round-numbered kink if workers also bunch at those for reasons beyond the jump in marginal tax rate. We control for this flexibly, including dummies for bins outside the bunching region that are multiples of $R = \{500, 1000\}$. This allows for the fact that some round numbers may be "rounder" than others and therefore induce more bunching. Furthermore, we include fixed effects for bins in the set of bracket thresholds $K = \{z_1^*, z_2^*, z_3^*\}$ whenever they lie outside the excluded region but are included in the estimation of the counterfactual, thus netting out any potential influence of nearby kinks from the counterfactual.

The estimated counterfactual distribution is the predicted bin count from equation (14), including the contributions of the round-number and nearby-kink dummies, but excluding the contribution of the excluded range, i.e.:

$$\hat{c}_j = \sum_{i=0}^{p} \hat{\beta}_i (z_j)^i + \sum_{r \in R} \hat{\rho}_r \mathbb{1}\Big[\frac{z_j}{r} \in \mathbb{N}\Big] + \sum_{k \in K} \hat{\theta}_k \mathbb{1}\Big[z_j \in K \wedge z_j \notin [z_L, z_U]\Big]$$

The bunching mass is the difference between the actual and counterfactual bin counts, given by:

$$B = \sum_{j=z_L}^{z_U} (c_j - \hat{c}_j)$$

---

[23] We use the R package **bunching** (Mavrokonstantis, 2019), available at `https://CRAN.R-project.org/package=bunching`, to obtain all bunching estimates in this paper.

To make responses comparable across different kinks, we normalize the bunching mass by the average counterfactual frequency in the excluded range, $\hat{c}_0$, to obtain the excess mass:

$$b = \frac{B}{\hat{c}_0}$$

where $\hat{c}_0 = \left[\frac{z_U - z_L}{\delta}\right]^{-1} \sum_{j=z_L}^{z_U} \hat{\beta}_i(z_j)$. We calculate bootstrapped standard errors using 500 bootstrap samples with replacement. The standard deviation of the distribution of these estimates yields the standard error of b.

Using $\hat{c}_0$, we can estimate $h_0(z^*) = \hat{c}_0/\delta$ and $\Delta z^* = B/h_0(z^*)$ which identifies the marginal bunching individual. The earnings response of the marginal buncher can then be used to estimate the observed elasticity of taxable income w.r.t. the net-of-tax rate:

$$\hat{\varepsilon} = \frac{\Delta z^*/z^*}{\Delta\tau/(1-\tau)}$$

To distinguish it from the structural elasticity estimates based on the models with adjustment frictions, we refer to this parameter as the observed elasticity calculated according to the Saez (2010) method throughout the paper.

Adjusting the Counterfactual Density. Some of the bunching literature additionally corrects the counterfactual density, shifting $\hat{c}_j$ upwards to the right of the kink. This is typically done to account for some unobserved counterfactual mass to the right of the kink when individuals bunch from above. In the baseline estimation, we do not perform such a correction for two reasons. First, in our setting, bunching may originate from above or from below (see Section 4) and thus it is not clear ex-ante how to adjust the density. Second, some of the literature argues against correcting the counterfactual density (e.g. Kleven 2016). Nevertheless, we estimate alternative corrected counterfactuals as robustness checks. On the one hand, we follow the procedure suggested by Chetty et al. 2011 to adjust the counterfactual density to the right of the kink, accounting for bunching originating from above as predicted by the standard model. On the other hand, we implement an analogous correction of the density to the left of the kink, accounting for bunching from below as predicted by the adapted model. As we discuss in Section 3.2, bunching estimates do not change substantially in both cases.

C.2 Identification of Static Models

The goal of estimating the static models is to obtain estimates of the elasticity ε and the fixed cost of adjusting earnings ϕ. Within both the standard and the adapted model, we derive equations linking bunching while a kink is in place and when it is removed to these parameters. These equations can then be estimated using the observed bunching moments in the data.

As is standard in the bunching literature, we assume iso-elastic, quasi-linear utility

$$u(c, z; n) = c(z) - \frac{n}{1+\frac{1}{\varepsilon}} \left(\frac{z}{n}\right)^{1+\frac{1}{\varepsilon}}$$

The first-order condition for optimal earnings supply is $z = n(1-\tau)^{\varepsilon}$.

C.2.1 Frictionless Bunching

Recall that bunching in the frictionless case is

$$B^* = \int_{z^*}^{z^*+\Delta z^*} h_0(z) dz \approx \Delta z^* \cdot h_0(z)$$

where the approximate equality holds if $h_0(z)$ is constant on $[z^*, z^* + \Delta z^*]$. There are two tangency conditions for the marginal buncher, $z^* = n(1-\tau_1)^{\varepsilon}$ and $z^* + \Delta z^* = n(1-\tau_0)^{\varepsilon}$. Combining the bunching equation with the tangency conditions yields

$$\frac{B^*}{h_0(z)} = z^* \left[\left(\frac{1-\tau_0}{1-\tau_1}\right)^{\varepsilon} - 1 \right] \tag{15}$$

Thus, bunching is only a function of the elasticity ε in the frictionless case.

C.2.2 Standard Model with Bunching from Above

Next, we turn to the identification of static models with adjustment frictions. First, consider the standard model of bunching with adjustment frictions following Gelber et al. (2020). When a kink is introduced, bunching under adjustment frictions is

$$B_1 = \int_{\underline{z}}^{z^*+\Delta z^*} h_0(z) dz \approx (z^* + \Delta z^* - \underline{z}) \cdot h_0(z)$$

There is no closed-form solution to connect observed bunching to the parameters of interest. Instead, the solution is defined by a combination of the bunching equation with (a) the tangency conditions of the upper marginal buncher at z^* and $z^* + \Delta z^*$, (b) the indifference condition of the lower marginal buncher from equation (3), and (c) the tangency condition of the lower marginal buncher at $\underline{z}$.

$$z^* + \Delta z^* = \left(\frac{1-\tau_0}{1-\tau_1}\right)^{\varepsilon} \cdot z^* \tag{16a}$$

$$c(\underline{z}) - \frac{\underline{n}}{1+\frac{1}{\varepsilon}} \left(\frac{\underline{z}}{\underline{n}}\right)^{1+\frac{1}{\varepsilon}} = c(z^*) - \frac{\underline{n}}{1+\frac{1}{\varepsilon}} \left(\frac{z^*}{\underline{n}}\right)^{1+\frac{1}{\varepsilon}} - \phi \tag{16b}$$

$$\underline{n} = \frac{\underline{z}}{(1-\tau_0)^{\varepsilon}} \tag{16c}$$

Equation (16) links observed bunching while a kink is in place to the two parameters ε and ϕ. The consumption level at $\underline{z}$ can be expressed as $c(\underline{z}) = c(z^*) + (\underline{z} - z^*)(1 - \tau_1)$.

When a kink is eliminated, remaining bunching with adjustment frictions is

$$B_0 = \int_{\underline{z}}^{\overline{z}} h_0(z)dz \approx (\overline{z} - \underline{z}) \cdot h_0(z)$$

Again, the solution is defined by a combination of the bunching equation with (a) the indifference condition of the marginal de-buncher from equation (5), (b) the tangency condition of the marginal de-buncher at $\overline{z}$, as well as the indifference and tangency conditions for the lower marginal buncher shown in equation (16).

$$c(z^*) - \frac{\overline{n}}{1+\frac{1}{\varepsilon}}\left(\frac{z^*}{\overline{n}}\right)^{1+\frac{1}{\varepsilon}} = c(\overline{z}) - \frac{\overline{n}}{1+\frac{1}{\varepsilon}}\left(\frac{\overline{z}}{\overline{n}}\right)^{1+\frac{1}{\varepsilon}} - \phi \tag{17a}$$

$$\overline{n} = \frac{\overline{z}}{(1-\tau_0)^{\varepsilon}} \tag{17b}$$

Equation (17) links observed bunching after a kink is removed to the two parameters ε and ϕ. The consumption level at $\overline{z}$ can be expressed as $c(\overline{z}) = c(z^*) + (\overline{z} - z^*)(1 - \tau_0)$.

C.2.3 Adapted Model with Bunching from Below

Second, we consider the adapted model where individuals start out from their optimal earnings level under a lower kink, such that bunching at the new kink originates from below. In this model, bunching at a kink that is introduced by shifting a bracket cutoff is

$$\tilde{B}_1 = \int_{z^*}^{\underline{\underline{z}}} h_0(z)dz \approx (\underline{\underline{z}} - z^*) \cdot h_0(z)$$

The solution is defined by a combination of the bunching equation with (a) the indifference condition of the marginal buncher from equation (7), and (b) the tangency conditions of the marginal buncher at $\underline{\underline{z}}$ and $\underline{\underline{z}}'$.

$$c(\underline{\underline{z}}') - \frac{\underline{\underline{n}}}{1+\frac{1}{\varepsilon}}\left(\frac{\underline{\underline{z}}'}{\underline{\underline{n}}}\right)^{1+\frac{1}{\varepsilon}} = c(z^*) - \frac{\underline{\underline{n}}}{1+\frac{1}{\varepsilon}}\left(\frac{z^*}{\underline{\underline{n}}}\right)^{1+\frac{1}{\varepsilon}} - \phi \tag{18a}$$

$$\underline{\underline{n}} = \frac{\underline{\underline{z}}}{(1-\tau_0)^{\varepsilon}} \tag{18b}$$

$$\underline{\underline{n}} = \frac{\underline{\underline{z}}'}{(1-\tau_1)^{\varepsilon}} \tag{18c}$$

Equation (18) links observed bunching while a kink is in place to the two parameters ε and ϕ. The consumption level at $\underline{\underline{z}}'$ can be expressed as $c(\underline{\underline{z}}') = c(z^*) - (z^* - \underline{\underline{z}}')(1 - \tau_0)$.

When the kink is removed, remaining bunching is

$$\tilde{B_0} = \int_{z^*}^{\overline{\overline{z}}} h_0(z)dz \approx (\overline{\overline{z}} - z^*) \cdot h_0(z)$$

The solution is defined by a combination of the bunching equation with (a) the indifference condition of the marginal de-buncher from equation (9), and (b) the tangency condition of the marginal de-buncher at $\overline{\overline{z}}$.

$$c(z^*) - \frac{\overline{\overline{n}}}{1+\frac{1}{\varepsilon}} \left(\frac{z^*}{\overline{\overline{n}}}\right)^{1+\frac{1}{\varepsilon}} = c(\overline{\overline{z}}) - \frac{\overline{\overline{n}}}{1+\frac{1}{\varepsilon}} \left(\frac{\overline{\overline{z}}}{\overline{\overline{n}}}\right)^{1+\frac{1}{\varepsilon}} - \phi \tag{19a}$$

$$\overline{\overline{n}} = \frac{\overline{\overline{z}}}{(1-\tau_0)^{\varepsilon}} \tag{19b}$$

Equation (19) links observed bunching after a kink is removed to the two parameters ε and ϕ. The consumption level at $\overline{\overline{z}}$ can be expressed as $c(\overline{\overline{z}}) = c(z^*) + (\overline{\overline{z}} - z^*)(1-\tau_0)$.

C.3 Identification of Dynamic Models

The goal of estimating the dynamic models is to obtain estimates of the elasticity ε, the fixed cost of adjusting earnings ϕ, and the probability of being subject to adjustment frictions in each period π_t. We can derive equations linking bunching at a kink in period t while it is in place or after it is removed to these parameters. In the data, we observe kinks in place for up to three years, and thus we estimate the cumulative probabilities π_1, $\pi_1\pi_2$ and $\pi_1\pi_2\pi_3$, in addition to ε and ϕ.

C.3.1 Standard Model with Bunching from Above

Bunching is given by equation (10) while the kink is in place. Hence,

$$\begin{aligned} B_1^1 &= \pi_1 \cdot B_1 + (1-\pi_1) \cdot B^* \\ B_1^2 &= \pi_1\pi_2 \cdot B_1 + (1-\pi_1\pi_2) \cdot B^* \\ B_1^3 &= \pi_1\pi_2\pi_3 \cdot B_1 + (1-\pi_1\pi_2\pi_3) \cdot B^* \end{aligned} \tag{20}$$

where B^* and B_1 are functions of parameters ε and ϕ as in equations (15) and (16). Hence, equation (20) links bunching observations B_1^t to parameters ε, ϕ and π_1, $\pi_1\pi_2$ and $\pi_1\pi_2\pi_3$.

After the kink is eliminated, remaining bunching is given by equation (11). Remaining bunching B_0^{T+t} now depends both on time since the kink was removed t and the duration for which the kink was in place T. The following example supposes the kink was in place for $T = 3$

periods:

$$\begin{aligned} T = 3: \quad B_0^{3+1} &= \pi_1 \cdot [B_0 + (1 - \pi_1\pi_2\pi_3) \cdot (B^* - B_1)] \\ B_0^{3+2} &= \pi_1\pi_2 \cdot [B_0 + (1 - \pi_1\pi_2\pi_3) \cdot (B^* - B_1)] \\ B_0^{3+3} &= \pi_1\pi_2\pi_3 \cdot [B_0 + (1 - \pi_1\pi_2\pi_3) \cdot (B^* - B_1)] \end{aligned} \tag{21}$$

where B^*, B_1 and B_0 are functions of parameters ε and ϕ as in equations (15), (16) and (17). Hence, equation (21) links bunching observations B_0^{T+t} to parameters ε, ϕ and π_1, $\pi_1\pi_2$ and $\pi_1\pi_2\pi_3$.

C.3.2 Adapted Model with Bunching from Below

Bunching is given by equation (12) while the kink is in place. Hence,

$$\begin{aligned} \tilde{B}_1^1 &= \pi_1 \cdot \tilde{B}_1 + (1 - \pi_1) \cdot B^* \\ \tilde{B}_1^2 &= \pi_1\pi_2 \cdot \tilde{B}_1 + (1 - \pi_1\pi_2) \cdot B^* \\ \tilde{B}_1^3 &= \pi_1\pi_2\pi_3 \cdot \tilde{B}_1 + (1 - \pi_1\pi_2\pi_3) \cdot B^* \end{aligned} \tag{22}$$

where B^* and $\tilde{B}_1$ are functions of parameters ε and ϕ as in equations (15) and (18). Hence, equation (22) links bunching observations $\tilde{B}_1^t$ to parameters ε, ϕ and π_1, $\pi_1\pi_2$ and $\pi_1\pi_2\pi_3$.

Remaining bunching after the kink is eliminated is given by equation (11). Hence,

$$\begin{aligned} \tilde{B}_0^{T+1} &= \pi_1 \cdot \tilde{B}_0 \\ \tilde{B}_0^{T+2} &= \pi_1\pi_2 \cdot \tilde{B}_0 \\ \tilde{B}_0^{T+3} &= \pi_1\pi_2\pi_3 \cdot \tilde{B}_0 \end{aligned} \tag{23}$$

where $\tilde{B}_0$ is a function of parameters ε and ϕ as in equation (19). Hence, equation (23) links bunching observations B_0^{T+t} to parameters ε, ϕ and π_1, $\pi_1\pi_2$ and $\pi_1\pi_2\pi_3$.

C.4 Allowing for Asymmetric Frictions

In some specifications, we allow for asymmetric frictions. In this case, we replace the single adjustment cost parameter ϕ by a bunching adjustment cost parameter ϕ_+ in equation (16) or equation (18). Analogously, we replace ϕ by a de-bunching adjustment cost parameter ϕ_- in equation (17) or (19). The static model can then be used to identify the parameters ε, ϕ_+ and ϕ_-. Taking these modified equations as an input into the dynamic equations (20) and (21) or (22) and (23), the dynamic model can be used to estimate ε, ϕ_+, ϕ_-, π_1, $\pi_1\pi_2$ and $\pi_1\pi_2\pi_3$.

D The Role of Income Effects

In this appendix, we discuss the potential role of income effects in the adapted model from Section 4.3. In Section D.1, we theoretically characterize bunching responses with income effects. In Section D.2, we present a simple calibration of income effects. We focus on the adapted model throughout, since income effects do not conceptually change the analysis in the standard model from Section 4.2.

D.1 Adapted Model with Income Effects

D.1.1 Introducing a Kink

First, we consider the situation where a new kink at z^* is introduced while an existing kink of the same size at $z_0^* < z^*$ is eliminated. In Figure D1, schedule (1) denotes the budget constraint under the old kink and schedule (0) is the budget constraint in the counterfactual scenario with linear net-of-tax rate $1-\tau_0$. Schedule (2) represents a hypothetical budget constraint with slope $1-\tau_1$ passing through the new kink at z^*. Potential bunching individuals A, B, C, D and E are located at earnings levels between z^* and $z^* + \Delta z^*$ under the counterfactual budget constraint (0) (points A_0 to E_0). All potential bunchers would locate at or below the kink z^* in the absence of income effects, which we denote by points A_2 to E_2 on schedule (2). However, in addition to the slope change, there is also a decrease in virtual income from y to y_0, corresponding to a parallel shift of the budget constraint from schedule (2) to schedule (1). Income effects induce individuals to slightly increase their earnings to points A_1 to E_1. The crucial difference to the simplified model without income effects is that some potential bunchers can now be located above z^* when the new kink is introduced. This can be seen by considering individual A who has counterfactual earnings $z^* + \Delta z^*$. A locates exactly at z^* under schedule (2), but locates somewhat above z^* under schedule (1). In the figure, the set of potential bunchers who are located above z^* under the old kink is given by individuals A to C, while potential bunchers C to E are still below z^*.

In the presence of income effects, there can be two potential marginal bunchers in Figure D1. First, individual D with ability $\underline{\underline{n}}$ who earns $\underline{\underline{z}}$ along schedule (0) (point D_0) is the *marginal buncher from below.* She changes her earnings to $\underline{\underline{z}}'$ in response to the old kink. When the new kink is introduced, she may decide to remain at earnings level $\underline{\underline{z}}'$ (point D_1'), or incur the adjustment cost and increase her earnings to bunch at z^*. All individuals with counterfactual earnings between z^* and $\underline{\underline{z}}$ bunch, as they moved the furthest below the new kink before it was introduced, and the utility difference between this earnings level and the kink makes it worthwhile incurring the adjustment cost. Second, individual B with ability n_B who earns $\underline{\underline{z}}_B$ along schedule (0) (point B_0) is the *marginal buncher from above.* She changes her earnings to $\underline{\underline{z}}_B'$ in response to the old kink. When the new kink is introduced, she may decide to remain at earnings level $\underline{\underline{z}}_B'$ (point B_1'), or incur the adjustment cost and decrease her earnings to bunch at z^*. All individuals with counterfactual earnings between $\underline{\underline{z}}_B$ and $z^* + \Delta z^*$ bunch, as they are located the furthest above the new kink before it was introduced. Individuals with counterfactual earnings between $\underline{\underline{z}}$ and $\underline{\underline{z}}_B$, who are closest to z^* under the old kink, do not bunch due to adjustment frictions.

The marginal buncher from below is characterized by the indifference condition

$$u(\underline{\underline{z}}', \tau_0; \underline{\underline{n}}) = u(z^*, \tau_0; \underline{\underline{n}}) - \phi \tag{24}$$

together with her tangency condition at $\underline{\underline{z}}$ under τ_0 and her tangency condition at $\underline{\underline{z}}'$ under τ_1. Note that these conditions are the same as those describing the marginal buncher in equation (7) from the simplified model without income effects.

The marginal buncher from above is characterized by the indifference condition

$$u(\underline{\underline{z}}_B, \tau_1; n_B) = u(z^*, \tau_0; n_B) - \phi \tag{25}$$

together with her tangency condition at $\underline{\underline{z}}_B$ under τ_0 and her tangency condition at $\underline{\underline{z}}'_B$ under τ_1.

Existence of Marginal Bunchers. To obtain predicted overall bunching, a crucial issue is whether both of these marginal bunchers exist. For the marginal buncher from below to exist, at least one individual must want to bunch from below. This is the case if individual E, who is the furthest below z^* when the new kink is introduced, wants to bunch:

$$u(\underline{\underline{z}}'_E, \tau_0; n_E) \leq u(z^*, \tau_0; n_E) - \phi \tag{26}$$

where ability n_E is pinned down by the tangency condition of individual E at z^* under τ_0 and her tangency condition at $\underline{\underline{z}}'_E$ under τ_1.

Similarly, at least one individual must want to bunch from above for the upper marginal buncher to exist. This is the case if individual A, who is the furthest above z^* when the new kink is introduced, wants to bunch:

$$u(\underline{\underline{z}}'_A, \tau_1; n_A) \leq u(z^*, \tau_0; n_A) - \phi \tag{27}$$

where ability n_A is pinned down by the tangency condition of individual A at $z^* + \Delta z^*$ under τ_0 and her tangency condition at $\underline{\underline{z}}'_A$ under τ_1.

Case 1: Bunching from Above and Below. If conditions (26) and (27) both hold, bunching occurs from below and from above, and the total bunching mass is given by

$$\tilde{B_1} = \int_{z^*}^{\underline{\underline{z}}} h_0(z)dz + \int_{\underline{\underline{z}}_B}^{z^*+\Delta z^*} h_0(z)dz \tag{28}$$

Case 2: Bunching Only from Below. If condition (26) holds but (27) does not hold, bunching occurs only from below:

$$\tilde{B}_1 = \int_{z^*}^{\underline{\underline{z}}} h_0(z)dz \tag{29}$$

Case 3: Bunching Only from Above. If condition (26) does not hold but (27) holds, bunching occurs only from above:

$$\tilde{B}_1 = \int_{\underline{\underline{z}}_B}^{z^*+\Delta z^*} h_0(z)dz \tag{30}$$

Case 4: No Bunching. If both conditions (26) and (27) do not hold, no bunching occurs at all.

Note that Case 2 is analogous to the simplified model without income effects from Section 4.3, since bunching is predicted to occur only from below and the marginal buncher is characterized by the same conditions.

D.1.2 Eliminating a Kink

Now suppose the kink is in turn eliminated. The marginal de-bunching individual can be found analogously to the simplified model without income effects. As shown in Panel (b) of Figure 7, the marginal de-buncher with ability $\bar{\bar{n}}$ earns $\bar{\bar{z}}$ under linear tax rate τ_0 and bunches z^* while the kink is in place. When the kink is subsequently removed, she is indifferent between remaining at z^* and incurring the adjustment cost to return to her counterfactual earnings level. The indifference condition of the marginal de-buncher is given by equation (5).

However, since the bunching mass while the kink is in place depends on conditions (26) and (27), the amount of residual bunching also differs across Cases 1 to 4. In addition, residual bunching can depend on the ordering of cutoffs $\bar{\bar{z}}$, $\underline{\underline{z}}$, and $\underline{\underline{z}}_B$.

Case 1: Bunching from Above and Below. If conditions (26) and (27) both hold, bunching occurred from below and above. Residual bunching depends on the ordering of cutoffs.

Case 1a: if $\bar{\bar{z}} \leq \underline{\underline{z}}$, remaining bunching at the kink is

$$\tilde{B}_0 = \int_{z^*}^{\bar{\bar{z}}} h_0(z)dz \tag{31}$$

Case 1b: if $\bar{\bar{z}} > \underline{\underline{z}}$ and $\bar{\bar{z}} \leq \underline{\underline{z}}_B$:

$$\tilde{B}_0 = \int_{z^*}^{\underline{\underline{z}}} h_0(z)dz \tag{32}$$

Case 1c: if $\overline{\overline{z}} > \underline{\underline{z}}_B$:

$$\tilde{B_0} = \int_{z^*}^{\underline{\underline{z}}} h_0(z)dz + \int_{\underline{\underline{z}}_B}^{\overline{\overline{z}}} h_0(z)dz \tag{33}$$

Case 2: Bunching Only from Below. If condition (26) holds but (27) does not hold, bunching occurred only from below. In this case, remaining bunching is

$$\tilde{B_0} = \int_{z^*}^{\overline{\overline{z}}} h_0(z)dz \tag{34}$$

Note that equation (34) implicitly assumes $\overline{\overline{z}} \leq \underline{\underline{z}}$. If this did not hold, there would be no de-bunching, i.e. residual bunching would be the same as bunching at the kink while it is in place, which would be inconsistent with the reduced-form empirical evidence.

Case 3: Bunching Only from Above. If condition (26) does not hold but (27) holds, bunching occurred only from above. Residual bunching is

$$\tilde{B_0} = \int_{\underline{\underline{z}}_B}^{\overline{\overline{z}}} h_0(z)dz \tag{35}$$

Note that equation (35) implicitly assumes $\overline{\overline{z}} > \underline{\underline{z}}_B$. If this condition did not hold, there would be no residual bunching at all, which would be inconsistent with the empirical evidence.

Case 4: No Bunching. If both conditions (26) and (27) do not hold, no bunching occurred. Thus, there is no residual bunching.

Again, we note that predicted de-bunching in Case 2 is analogous to the simplified model without income effects from Section 4.3.

D.2 Calibrating Income Effects

In Section 3, we document empirically that bunching builds gradually when kinks are newly introduced, and bunching does not immediately disappear when kinks are removed. In principle, these empirical patterns could be rationalized by any of the theoretical Cases 1 to 3 above, as all of these predict some (but incomplete) bunching while a kink is in place and some residual bunching when it is removed.[24] Which of the cases applies ultimately depends on the magnitude of income effects. If income effects are small or zero, potential bunchers from above are located too close to the kink, such that they do not find it worthwhile to incur adjustment costs to bunch from above, and bunching occurs only from below as in Case 2. If income effects were

---

[24] We also observe no bunching at some kinks as in theoretical Case 4, but these kinks are not used in the structural estimation.

large, on the other hand, bunching could occur from two disjoint segments below and above the kink (Case 1) or, in extreme cases, only from above (Case 3). To make progress and gauge how income effects impact bunching empirically, we next calibrate their potential magnitude.

As in Section 4.1, we can write optimal earnings as a function $z = z(1-\tau, y; n)$ where y is virtual income in the tax bracket where the agent is located.[25] Earnings responses to tax reforms can be decomposed via the Slutksy equation

$$\frac{dz}{d(1-\tau)} = \frac{dz^c}{d(1-\tau)} + \frac{dz}{dy}z$$

where z^c is the compensated earnings supply function. Thus, total earnings responses equal the sum of the compensated earnings responses and income effects. Income effects can be represented as earnings responses to changes in virtual income, or the "marginal propensity to earn" dz/dy.

In our empirical setting, kinks are introduced by tax reforms moving existing kinks upwards. As Figure D1 shows, such tax reforms induce an increase in virtual income corresponding to a parallel upward shift of the budget constraint. Virtual income in tax bracket i can be calculated as

$$y_i = y_{i-1} + \left[(1-\tau_{i-1}) - (1-\tau_i)\right] k_i \tag{36}$$

where k_i is the lower cutoff of the tax bracket. For instance, at the CYP 10,000 kink, virtual income under the old kink is $d\tau \cdot z_0^* = 0.2 \cdot 9000 = 1800$, virtual income under the new kink is $d\tau \cdot z^* = 0.2 \cdot 10000 = 2000$, and thus the change in virtual income is $dy =$CYP 200.

In order to calibrate potential earnings responses to such changes in virtual income, we need an estimate of the marginal propensity to earn (MPE). The literature offers some estimates of the MPE, including from studies on negative income tax experiments (e.g. Burtless 1986, Ashenfelter and Plant 1990), unconditional cash transfers (e.g. Akee et al. 2010, Jones and Marinescu 2019), and lottery winners (Imbens et al. 2001, Cesarini et al. 2017). These studies tend to find a small MPE between zero and –0.1. In addition, there is some evidence from social security settings suggesting a larger MPE of around –0.5 to –1 (e.g. Gelber et al. 2017, Giupponi 2019). These larger estimates reflect long-run responses and are largely driven by labor force participation, such that they are arguably less comparable to our setting. Nevertheless, we consider a wide range of assumptions about the value of the MPE in our calibrations below.

Ultimately, our goal is to understand whether income effects are large enough to induce bunching from above, which would change the analysis relative to the simplified model. We argue that comparing the magnitude of income effects to observed bunching responses can be informative for this purpose. Recall that some bunching from above is predicted to occur if individual A's earnings under the old kink $\underline{\underline{z}}'_A$ are sufficiently far above z^* to make it worthwhile

---

[25] The exposition here follows Bastani and Selin (2014) who incorporate income effects into bunching estimation. Like them, we abstract from any non-labor income, such that virtual income is zero in the first tax bracket. This simplification is closely in line with our main sample definition, where we focus on wage earners who only have earnings from employment.

to incur adjustment costs and bunch. In order to gauge whether individual A is likely to bunch, we would need to know what distance from the kink is sufficient to make incurring adjustment costs worthwhile. For this purpose, we can compare A's distance from the kink to that of the lower marginal buncher (individual D), the individual closest to the kink known to bunch.[26]

To calculate the relevant distances from the kink, we can use calibrated income effects in combination with the information contained in our bunching estimates. The distance between $\underline{\underline{z}}'_A$ and z^* corresponds to the income effect induced by the change in virtual income when the kink is introduced. The total segment of potential bunchers, that is the distance between $\underline{\underline{z}}'_E$ and $\underline{\underline{z}}'_A$, approximately equals Δz^*, the width of the frictionless bunching segment. The segment containing the subset of bunchers from below is $\underline{\underline{z}}'_E$ to $\underline{\underline{z}}'$. Thus, the distance between the lower marginal buncher and the kink $(z^* - \underline{\underline{z}}')$ can be calculated by subtracting the width of the bunching segment from below $(\underline{\underline{z}}' - \underline{\underline{z}}'_E)$ and the income effect $(\underline{\underline{z}}'_A - z^*)$ from Δz^*.

Table D1 quantifies this intuition in simple calibrations. For each kink used in the structural estimation, column (2) shows the change in virtual income Δy when it is introduced, calculated according to equation (36). The increases in virtual income are modest, corresponding to at most 2% of the income level at the kink.[27] Column (3) shows the width of the full potential bunching segment Δz^*. We calculate Δz^* based on hypothetical frictionless earnings responses to each kink implied by the structural elasticity estimates from Panel C of Table 2. Column (4) shows the width of the bunching segment from below $\Delta\underline{\underline{z}} \equiv \underline{\underline{z}}' - \underline{\underline{z}}'_E$, which we calculate based on the estimated excess mass while the respective kink is in place. Columns (5), (7), (9) and (11) then show earnings responses due to income effects $IE \equiv \underline{\underline{z}}'_A - z^*$, which we calibrate for a wide range of MPE values between –0.1 and –1. Finally, columns (6), (8), (10) and (12) show the implied distance between the lower marginal buncher and the kink $z^* - \underline{\underline{z}}'$ under each scenario.[28]

The key comparison in the table is the distance of individual A from the kink given by IE vs. the distance of the lower marginal buncher from the kink $z^* - \underline{\underline{z}}'$. As discussed above, the MPE values most directly applicable to our setting are –0.1 or even smaller. At the CYP 10,000 kink, a MPE of –0.1 implies that individual A would be located CYP 20 above the kink, while the lower marginal buncher is CYP 487.8 below the kink. Thus, individuals located less than CYP 487.8 below the kink do not bunch due to adjustment costs. This makes it very unlikely that individual A, who is only CYP 20 above the kink, would find it worthwhile to incur adjustment

---

[26] This comparison requires that the benefit of relocating to the kink is increasing in the distance to the kink, which we assume throughout the theoretical analysis.

[27] In fact, there are two cases of kinks where virtual income does not increase upon their introduction. At the CYP 9,000 kink virtual income does not change, and at the CYP 12,000 kink virtual income decreases in comparison to its prior location. This occurs due to additional changes in marginal tax rates before these kinks were introduced. Hence, there are no income effects at the CYP 9,000 kink and, if anything, income effects would push potential bunchers further below the CYP 12,000 kink.

[28] Since total observed bunching could include some bunching from above, the distance $\Delta\underline{\underline{z}}$ we calculate in column (4) provides an upper bound on the segment of bunchers from below. Thus, the distances shown in columns (6), (8), (10) and (12) correspond to lower bounds on $z^* - \underline{\underline{z}}'$ under the respective MPE value. If anything, this strengthens our conclusion that the lower marginal buncher is much further from the kink than any potential buncher from above.

costs to bunch. Similarly, we find that individual A is much closer to the kink than the lower marginal buncher across all kinks and all values of the MPE we consider. Even under a MPE of –1, which is much larger than comparable estimates from the literature, the calibration implies that the lower marginal buncher is at least 50% further from the kink than any potential buncher from above. Thus, even for large values of the MPE, bunching from above due to income effects is unlikely to occur.[29]

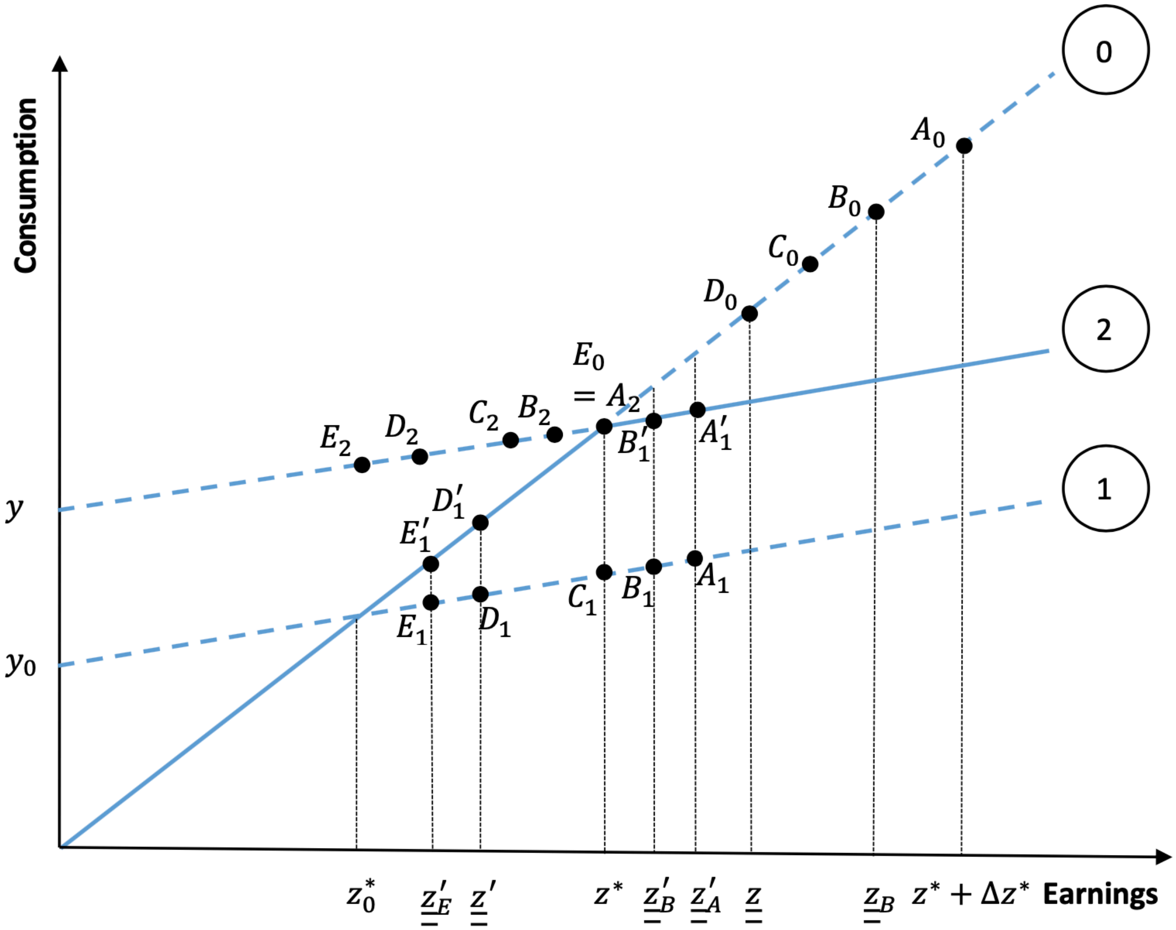
We conclude that income effects are unlikely to fundamentally change the analysis in our setting. Since income effects calibrated under a range of MPE values do not lead to bunching from above, bunching at new kinks occurs only from below under the adapted model. This motivates our use of the simplified model without income effects from Section 4.3, which yields analogous bunching predictions to the full model when income effects are small.

Finally, it is worth clarifying that the results here do not imply that bunching from above cannot occur at all, but that income effects are unlikely to create an additional bunching segment to the right of the kink. Bunching from above can still occur if some individuals start out from their counterfactual earnings level above the kink, as is the case in the standard model from Section 4.2.

---

[29] An additional implication of the calibration results presented here is that marginal bunchers always start from an interior solution, as assumed in Section 4.3. To see this, note that Δz^* in Table D1 is always smaller than the distance between old and new kinks shown in Figure 1.

Figure D1: Frictions and Bunching in the Adapted Model with Income Effects



Notes: The figure illustrates how frictions attenuate bunching responses in the adapted model with income effects from Section D.1. Schedule (0) denotes the budget constraint with linear tax rate τ_0, schedule (1) is the budget constraint under the old kink at z_0^*, and schedule (2) is a hypothetical budget constraint with tax rate τ_1 passing through the new kink at z^*. Points A_i to E_i denote the optimal earnings choice of individuals A to E under the respective schedule i. The key difference to the simplified model without income effects is that some potential bunchers (individuals A to C) are located above z^* when the new kink is introduced, while other potential bunchers (individuals C to E) are still located below z^*. Individual D with counterfactual earnings $\underline{\underline{z}}$ is the marginal buncher from below, who is indifferent between remaining at $\underline{\underline{z}}'$ and bunching at z^*. Individual B with counterfactual earnings $\underline{\underline{z}}_B$ is the marginal buncher from above, who is indifferent between remaining at $\underline{\underline{z}}'_B$ and bunching at z^*.

Table D1: Calibrating Income Effects

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
				MPE=–0.1		MPE=–0.3		MPE=–0.5		MPE=–1	
Kink (z^*)	Δy	Δz^*	$\Delta \underline{\underline{z}}$	IE	$z^* - \underline{\underline{z}}'$	IE	$z^* - \underline{\underline{z}}'$	IE	$z^* - \underline{\underline{z}}'$	IE	$z^* - \underline{\underline{z}}'$
CYP 10,000	200	574	66.2	20	487.8	60	447.8	100	407.8	200	307.8
CYP 9,000	0	450	15.7	0	434.3	0	434.3	0	434.3	0	434.3
CYP 12,000	-500	187.5	52.1	-50	185.5	-150	285.5	-250	385.5	-500	635.5
CYP 10,750	150	537.5	72.9	15	449.6	45	419.6	75	389.6	150	314.6
CYP 11,413	132.6	570.7	105.9	13.3	451.5	39.8	424.9	66.3	398.4	132.6	332.1

Notes: The table shows results from the calibration exercises described in Section D.2. For each of the kinks listed in column (1), the table shows the change in virtual income upon its introduction Δy in column (2), the total width of the potential bunching segment Δz^* in column (3) and the width of the bunching segment from below $\Delta \underline{\underline{z}}$ in column (4). Columns (5) to (12) present calibrated income effects IE and the distance of the marginal buncher from below from the kink $z^* - \underline{\underline{z}}'$ under a range of assumptions about the marginal propensity to earn (MPE).