



# A data-driven procedure to determine the bunching window: an application to the Netherlands

Nicole Bosch<sup>1,2</sup>  Vincent Dekker<sup>3</sup> · Kristina Strohmaier<sup>4</sup>

Published online: 3 February 2020  
© The Author(s) 2020

## Abstract

We extend the bunching approach introduced by Saez (Am Econ J Econ Policy 2:180–212, 2010) by proposing an intuitive, data-driven procedure to determine the bunching window. By choosing the bunching window ad hoc, researchers throw away informative data points for estimating the counterfactual income distribution in the absence of the kink. Assuming a descending bunching mass to both sides of the threshold, the proposed algorithm produces a distribution of lower and upper bounds for the bunching window. In each iteration, the bunching window is defined as all contiguous bin midpoints around the threshold that lie outside of the confidence band resulting from running a local regression through all data points outside of the excluded region. Monte Carlo simulations provide evidence that our data-driven procedure outperforms larger bunching windows in terms of bias and efficiency. In our application for the Netherlands, we find clear evidence of bunching behaviour at all three thresholds of the Dutch tax schedule with a precisely estimated elasticity of 0.023 at the upper threshold, which is driven by self-employed, women and joint tax filers.

**Keywords** Bunching · Elasticity of taxable income · Optimal bunching window

**JEL Classification** H21 · H24

---

✉ Kristina Strohmaier  
k.strohmaier@uni-tuebingen.de

<sup>1</sup> CPB Netherlands Bureau for Economic Policy Analysis, Bezuidenhoutseweg 30, 2594 AV The Hague, The Netherlands

<sup>2</sup> VU University Amsterdam, Amsterdam, The Netherlands

<sup>3</sup> Department of Economics, Chair of Public Economics, University of Hohenheim, Schloss Hohenheim 1D, 70599 Stuttgart, Germany

<sup>4</sup> Department of Economics, Chair of Public Economics, University of Tübingen, Nauklerstraße 47, 72074 Tübingen, Germany

## 1 Introduction

A central topic in public economics is the assessment of welfare losses caused by behavioural responses to income taxation. Following the seminal paper by Feldstein (1995), a large study emerged where welfare losses are inferred from the elasticity of taxable income (ETI).<sup>1</sup> Notwithstanding the large variation in identification strategies and data used in these studies, a common finding is that the elasticities and thus the tax-induced welfare losses are modest. Recent studies hint at different explanations for these modest estimates, such as optimisation frictions (Bastani and Selin 2014; Chetty et al. 2011), shifting of income over time (Le Maire and Schjerning 2013) or shifting across tax bases (Harju and Matikka 2016). More fundamentally, other papers claim that the structural parameter cannot be retrieved from these estimates, because the ETI depends on the institutional framework, such as the exact definition of taxable income (Slemrod 1998; Saez et al. 2012; Doerrenberg et al. 2017).

A growing strand of the literature utilises the bunching method to obtain a non-parametric estimate of the ETI (Saez 2010; Chetty et al. 2011). This method exploits the clustering behaviour of individuals at kinks in a nonlinear tax system to identify the ETI by the number of individuals that adjust their income to stay below the threshold of a tax bracket. Using the bunching method is attractive as it builds on a sound theoretical foundation and is not susceptible to endogeneity biases, a problem suffered by the previous ETI literature (Saez 2010; Gruber and Saez 2002; Weber 2016).<sup>2</sup>

The aim of our study is twofold: first, we tackle the issue of selecting the “bunching window”, which is necessary as individuals are not able to perfectly adjust their taxable income to the tax threshold.<sup>3</sup> The large number of robustness checks in previous studies already hints at the uncertainty regarding the optimal choice of the bunching window and the appropriate counterfactual model. In his seminal work, Saez (2010) already points out the importance of selecting the bunching window (which is labelled bandwidth in his application). On p. 196, he states that “[...] *in the presence of clustering around a kink point instead of exact bunching, the choice of the bandwidth matters. In this paper, we use a simple graphical visual approach for selecting the bandwidth which is a significant limitation. As mentioned above, with larger datasets and hence smoother density estimates, it could be possible to devise a method to detect bunching humps statistically and hence choose the bandwidth  $\delta$  with a systematic econometric method*”. The selection of the bunching window is

<sup>1</sup> See Saez et al. (2012) and Kleven (2016) for a comprehensive overview.

<sup>2</sup> Please note that this method hinges on the assumption that the derivative of the counterfactual density function  $h_0(z)$  with respect to  $z$  is continuous in  $z$  for all  $z$ . We thereby implicitly assume that the distribution of preference heterogeneity is restricted around the kink; otherwise, the bunching estimator would not be identified. See Blomquist and Newey (2017) for more details.

<sup>3</sup> Saez (2010), in his seminal bunching paper, mentions three reasons why we observe a bunching window instead of a clear spike. First, individuals may not be able to perfectly control their taxable income. Second, they may be unable to predict their income correctly, and third, they may have imperfect information.

thereby crucial in terms of both efficiency and bias. As for fear of (i) omitting bunching individuals and/or (ii) including bunchers during the estimation of the counterfactual distribution, researchers are often conservative in choosing the bunching window and thus throw away informative data points by using a too large bunching window which is inefficient.<sup>4</sup> However, even more important, the unbiasedness of the ETI depends on the bunching window in two ways as it does not only enter the estimation of the excess mass of individuals around the threshold (by setting the outer boundaries) but also directly affects the estimation of the counterfactual density following Chetty et al. (2011) that is needed to derive the elasticity. We therefore propose a simple, data-driven procedure to determine the bunching window. This improves on the visual inspection that determines the bunching window in the previous literature. As a consequence, our method does not impose any restrictions on the (a)symmetry of the bunching window around the threshold which has been shown to be important in the literature (Devereux et al. 2014; Lediga et al. 2018). Furthermore, given that there seems to exist a publication bias in the ETI literature (Neisser 2017), an automated, data-driven procedure for at least one of the parameters within the estimation procedure seems to be a valuable contribution on its own.

The basic idea of our data-driven method to determine the bunching window is simple and follows a straightforward stepwise procedure. Assuming a uni-modal or equal distribution of optimisation frictions, we expect the bunching mass to be descending to both sides of the threshold. In a first step, one arbitrarily chooses an excluded region, runs a local polynomial regression through the remaining data points and predicts the frequencies including corresponding confidence intervals. We then define the bunching window as all contiguous bin midpoints around the threshold that lie outside of the confidence band. In a second step, by iterating through all possible combinations of excluded regions, one obtains a distribution of lower and upper bounds of the bunching window.

In the second part of the study, we estimate the compensated elasticity of taxable income with respect to the net-of-tax rate in the Netherlands using the refined bunching approach. We employ a unique longitudinal data set containing exact declared taxable income and tax deductions for a representative sample of the Dutch population (IPO data from 2003 to 2014). Information on taxable income and deductions is provided by the Dutch tax authority and, therefore, free of measurement errors—something that is vital to obtain reliable estimates with the bunching method. The data also contain covariates, such as gender and marital status as well as information on self-employment, which enable us to analyse various sub-samples.

---

<sup>4</sup> Bastani and Selin (2014), for example, explicitly state in their paper that they employ a “wide bunching window” in a first step. Comparing the chosen bunching window (“a window of DKr 15,000 centred around the kink”) in Chetty et al. (2011) with the graphical bunching results, the bunching window has to be chosen ad hoc, because (1) the excess mass is not centred around the threshold and (2) the window of (− 7,7) seems to be too large. While Devereux et al. (2014) use an asymmetric window by eyeballing, the window seems also to be larger than necessary (at least for Panels A, B and D in Fig. 1, p. 38). Depending on the bunching window used, the point estimate for the ETI in their study varies between 0.10 and 0.14, which is, in our view, a non-negligible difference.

In addition, we are able to analyse the anatomy of responses using information on mortgage interest deductions.

Our main findings are as follows. First, Monte Carlo simulations show that our data-driven procedure delivers unbiased ETI estimates which are more precisely estimated compared to alternative bunching windows (i.e. show the smallest standard error). This key result is, in addition, robust to various changes in key parameters, such as different binwidths, degrees of optimisation frictions and functional form assumptions. The superiority of the approach is a valuable addition to the literature that has previously relied on visual inspection. Second, in our empirical application, we estimate an ETI with respect to the net-of-tax rate of 0.023 at the highest tax threshold which is highly significant. This result is in line with some of the bunching literature, such as in Chetty et al. (2011), who find an elasticity below 0.02 for their full sample of Denmark, but differs from, for example, Bastani and Selin (2014) who report an elasticity of only 0.004 for Swedish taxpayers. Third, we find significantly higher compensated ETIs for women and self-employed individuals. However, contrary to most other studies, we are also able to identify a nonzero elasticity for individuals in paid employment. Fourth, by analysing the anatomy of response, we find that most employees reduce their taxable income by utilising mortgage interest deductions that can be shifted between joint filers. Since bunching is absent among single tax filers who do not have this shifting possibility, we conclude that income shifting drives the result and that real responses (hours or investment decisions) to taxation are modest. Our descriptive finding that hourly wages between bunchers and non-bunchers do not differ corroborates this conclusion.

The paper proceeds with Sect. 2, which explains our data-driven procedure to determine the bunching window. Subsequently, the institutional setting and the data are presented in Sects. 3 and 4, respectively. Section 5 presents our estimation results. Section 6 provides the conclusion.

## 2 Methodological extension on the bunching window

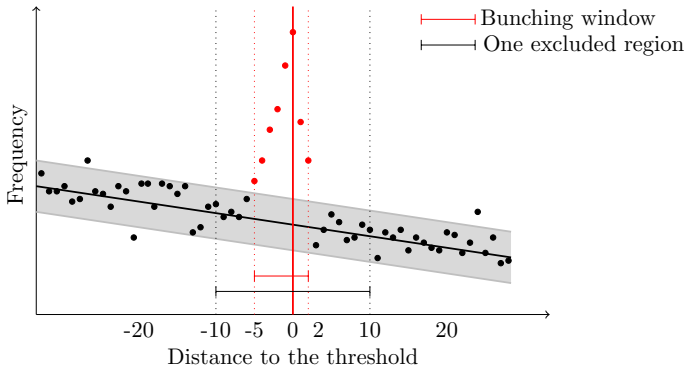
A major drawback of the nowadays well-established bunching method<sup>5</sup> is its sensitivity to the choice of several parameters, including the choice of the bunching window (Adam et al. 2015). The common practice in the literature is to select the window by visual inspection, which makes it vulnerable as it is selected at the researcher's discretion.

To the best of our knowledge, we are the first to extend the common practice within the bunching estimation by relying on the data at hand rather than on visual inspection to determine the bunching window.<sup>6</sup> Given the existing publication bias

---

<sup>5</sup> A short description of the bunching method is given in "Bunching method" section of Appendix 1.

<sup>6</sup> The study by Kleven and Waseem (2013) is one notable exception. Exploiting notches instead of kinks, the authors emphasise the important role of the bunching window. These authors also propose to use an iterative procedure, but only for the upper bound of the window. As their identification strategy exploits notches, they use the "missing mass" as a source to identify this upper bound, which is impossible for the case of kinks.



**Fig. 1** Data-driven procedure to determine the bunching window. *Notes* This figure shows the bin mid-points as well as the fitted values of a linear regression. The grey confidence band is calculated with the standard errors of the point prediction. Here, eight contiguous bin midpoints around the threshold lie outside the confidence band and therefore determine the relevant (asymmetric) bunching window, indicated by the red dotted line. An exemplary one of the excluded regions that we iterate over is shown by the outer grey dotted vertical lines (Color figure online)

within the ETI literature (Neisser 2017), removing the researcher’s discretion in this matter is preferable in its own right, but we can also show that our method produces elasticity estimates not only which are closer to the true elasticity, but which are more precisely estimated. In the following, we define the “optimal” bunching window to comprise all income bins that contain individuals who would adjust their taxable income as a response to the tax change at the threshold (so-called bunchers). If the bunching window is set too small, this directly affects the estimation of the ETI as it omits bunchers and thus underestimates the response. If the bunching window is set too large, the researcher throws away very informative data points during the estimation of the counterfactual density.<sup>7</sup> In addition, the existing literature typically implements symmetric bunching windows around the kink with varying sizes (again determined by graphical inspection). Our proposed method does not impose any restrictions on the (a)symmetry of the window as it is purely driven by the data at hand.

To determine the bunching window using binned data, we propose the following stepwise procedure:

1. Start with an excluded region of  $(x_-, x_+) = (0, 0)$  around the threshold, i.e. only omitting the bin where the threshold is located.
2. Run a local regression through all observed data bins outside the excluded region and predict the frequencies.

<sup>7</sup> This is (1) not efficient and (2) even more problematic if the shape of the income distribution is non-linear which leads to the underestimation (in case of a concave function) or overestimation (in case of a convex function) of the counterfactual density. This has already been noticed by Saez (2010) in his notes to Fig. 2.

3. Compute a confidence interval around the prediction.
4. Contiguous bin midpoints around the threshold bin that lie outside the confidence interval comprise the bunching window where the most left (right) bin point becomes the lower (higher) bound of the bunching window for this excluded region.
5. Repeat steps (2) to (4) for all other possible excluded regions  $(x_-, x_+)$  with  $x_- \in \{-X, (-X + 1), \dots, 0\}$  and  $x_+ \in \{0, 1, \dots, X\}$  to obtain a distribution of lower and upper bounds of the bunching window.
6. The mode of each distribution delivers the final bunching window comprised of a lower bound  $l^*$  and an upper bound  $u^*$ .

The formal derivation of this procedure is found in “Formal derivation bunching window” section of Appendix 1. A graphical intuition of our method for the simple case of a linear income distribution is given in Fig. 1.<sup>8</sup> The figure plots taxable income bins (relative to the threshold) against the number of individuals per income bin for an exemplary excluded region going from  $x_- = -10$  to  $x_+ = 10$ . All contiguous binpoints around the threshold that have a higher actual number of taxpayers than predicted (which are outside the confidence band, in red) comprise the bunching window, here going from  $-5$  to  $2$ . This procedure is repeated for each and every excluded region  $(x_-, x_+)$  producing  $(X + 1)^2$  estimates of  $l$  and  $u$  where  $X$  denotes the last bin used as excluded region.<sup>9</sup>

As we propose to iterate through all possible combinations of upper and lower bounds of the excluded region, this hedges against concerns that the chosen excluded region could affect the determination of the bunching window. It is important to note that in the end,  $X$  (i.e. the last bin used as excluded region) only determines the number of iterations to obtain the distribution of lower and upper bounds of the bunching window. If  $X$  is, for example, set to  $X = 20$ , steps (2) to (4) have to be repeated  $(X + 1)^2 = 441$  times which will lead to 441 observations for each of the distributions.<sup>10</sup> The confidence band should be set using standard significance levels, with higher levels tending to lead to a smaller bunching window. The benefit of shifting the point of arbitrariness away from the bunching window, which is used in the actual estimation of the elasticity, to the excluded region is that through iterations over several excluded regions, the dependency of the result upon the arbitrarily chosen parameter diminishes. Furthermore, the statistic of interest is no longer directly influenced by an arbitrarily chosen parameter.

<sup>8</sup> For illustrative purposes, we depict a linear counterfactual model. In our empirical application, the counterfactual is allowed to be a higher-order polynomial model, determined by the minimum BIC value.

<sup>9</sup> The size of  $X$  has to be chosen by the researcher ex ante. It might depend on the binsize, the tax schedule (e.g. the distance to other thresholds) and the respective income distribution. In practice, this is also related to the estimation sample in previous bunching studies to run the local regression.

<sup>10</sup> In the interest of computational time, we have restricted our simulation to the testing region from  $-15$  to  $15$  as our results indicated virtually no sensitivity of the final bunching window  $(l^*, u^*)$  if we further increased the size of the excluded region. Stated differently, we set  $X = 15$  and run  $(15 + 1)^2 = 256$  regressions per analysis.

The parameters for the optimal bunching window can then be used in the standard estimation technique as developed in Chetty et al. (2011). To estimate the ETI (see Eq. 4 in “Bunching method” section of Appendix 1), the excess mass  $b$  is the only parameter that needs to be estimated, as the other parameters are known policy parameters:

$$\hat{b} = \frac{\hat{B}}{\frac{\sum_{j=l^*}^{u^*} \hat{N}_j}{(u^* - l^* + 1)}} \quad \text{with} \quad \hat{B} = \sum_{j=l^*}^{u^*} (N_j - \hat{N}_j), \tag{1}$$

where  $\hat{B}$  denotes the number of bunchers within the bunching window.  $\hat{N}_j$  represents the counterfactual number of individuals within an income bin  $j$  that are determined by local polynomial regression of the form:

$$\hat{N}_j = \sum_{i=0}^q \beta_i \cdot X_j^i + \sum_{s=l^*}^{u^*} \gamma_s \cdot I[X_j = s] + \varepsilon_j. \tag{2}$$

In essence, the bunching method relates the number of bunching individuals (numerator of Eq. (1)) to the number of individuals who should be at the threshold in the absence of the tax kink (which is the average counterfactual number of individuals within the bunching window per bin, see denominator of Eq. (1)). By looking at Eqs. (1) and (2), one can see that the bunching window has a crucial role in determining the excess mass  $b$ . Potential bias from setting a wrong bunching window might arise from three distinct sources. First, setting a wrong bunching window directly affects the number of bunchers by summing over a wrong number of income bins to count them (summation bounds in the numerator of Eq. (1)). Trivially, if the bunching window is set too small, this would underestimate the excess mass and therefore the elasticity. A naive approach would be to be cautious by setting the bunching window relatively large. However, this would neglect that the number of bunchers  $\hat{B}$  is also affected through the estimation of the counterfactual density. So a second source of bias might arise from a too low or too high counterfactual, again, entering the calculation of  $\hat{B}$  (see  $\hat{N}_j$  within the summation to get  $\hat{B}$  which also depends on the bunching window). Finally, Eq. (1) makes clear that the counterfactual density also enters the denominator, which might be a third bias. Summing up, it is important to note that bounds of the bunching window are a central ingredient within the bunching estimation.

Next to the data-driven procedure for the determination of the bunching window, we rely on the Freedman–Diaconis rule to determine the bin size in each estimation. It states that the optimal binwidth is given by:  $2 \cdot IQR(x) \cdot n^{-\frac{1}{3}}$ , where IQR stands for interquartile range (Freedman and Diaconis 1981). Prior works have chosen a single binwidth for their analysis and subsequently altered the binwidth in robustness checks to show that the estimates are robust to the choice of the binwidth. The size of the binwidth, however, should depend on the number of observations around the threshold and therefore be determined for each threshold separately. In addition, we use the BIC criterion to determine the optimal number of polynomials when running the local regression.

## 2.1 Evaluation

In this section, we start by evaluating the performance of our endogenously determined bunching window by Monte Carlo simulations.<sup>11</sup> We will, however, also assess the real-world performance of our proposed method within our application (see especially Sect. 5.1).

In our Monte Carlo simulation, we use a generalised version of the standard labour supply model where the individual does not choose his labour supply, but taxable income. The individual will supply taxable income until the marginal utility of leisure (or marginal disutility of supplying taxable income) equals the marginal net-of-tax rate. The supply function is derived from a quasi-linear and iso-elastic utility function as in Saez (2010):

$$u(c, z) = c - \frac{z_0}{1 + \frac{1}{\epsilon}} \left( \frac{z}{z_0} \right)^{1 + \frac{1}{\epsilon}}.$$

Maximising this utility function with respect to the budget constraint  $c = (1 - t)z + R$  leads to the simple income supply function  $z = z_0(1 - \tau)^\epsilon$  where  $z$  is taxable income,  $z_0$  denotes “potential income”,  $\tau$  is the marginal tax rate and  $\epsilon$  the ETI. In the baseline simulation, we start with a triangular income distribution, where the functional form is known to be linear. To be more precise, the distribution of incomes  $z_0$  is drawn from the following triangular distribution:

$$z_0 = 80,000 - 60,000\sqrt{(1 - u)},$$

where  $u$  is drawn from a random uniform distribution. Our simulation with 600 repetitions draws  $N = 800,000$  incomes from a range between 20,000 and 80,000, with a threshold  $k$  at  $z = 40,000$ , a binwidth of 100 and a tax change of 10 percentage points. To model optimisation frictions, we implement additive optimisation frictions of 1% of  $k$ , which are distributed across the bunchers following a positively skewed beta distribution which—by construction—will lead to more mass to the left of the threshold. Observed taxable income choices are thus given by

$$z^{\text{observed}} = z^* + 0.01k(\mu - 0.8) \quad \text{with} \quad \mu \sim \text{Beta}(5, 2).$$

Please note that any uni-modal or equal distribution of optimisation frictions would result in a situation where the bunching mass is finally descending to both sides of the threshold (which is what we observe in all bunching studies). Given that there is typically only one single threshold relevant for the respective individual

<sup>11</sup> Our preferred method of validation for the endogenous procedure to determine the bunching window would be to replicate previous studies that detected the bunching window by eyeballing. In the taxation literature, however, most bunching studies rely on administrative data sets of personal taxable income, which are not freely available. For example, Danish micro data as used by Chetty et al. (2011) can only be accessed through a Danish partner institution and the restricted PSID files used by Saez (2010) significantly differ from the public use files and are hard to obtain outside the USA.



(who tries to reduce his income up to this point), assuming these kinds of optimisation frictions seems plausible to us.

We conduct our simulations for four true elasticities:  $e = 0.02$ ,  $e = 0.05$ ,  $e = 0.1$  and  $e = 0.5$  which resemble in their magnitude the elasticities found in the previous ETI literature. Based on the form of optimisation frictions assumed, we will not compare our procedure to the usually implemented symmetric bunching windows, but rather to a specification, where the bunching window is two (four) binpoints larger on both sides.<sup>12</sup>

Figure 2 shows the distribution of the ETI for the four true elasticities (indicated by the vertical line in each graph)  $e = 0.02$ ,  $e = 0.05$ ,  $e = 0.1$  and  $e = 0.5$  using the 600 replications of our baseline simulation. The solid line represents the results when using our data-driven procedure, while the dashed and dotted lines show the results where two and four bins were added on each side, respectively.

Starting with Fig. 2a, one can see that all three bunching windows (data-driven, data-driven+2 bins, data-driven+4 bins) deliver elasticity distributions which are centred around the true elasticity of  $e = 0.02$ . However, it is already clear from the graph that the variance of the elasticity estimate is smallest when using our data-driven procedure. This result can also be confirmed for the other cases (b) to (d) with larger true elasticities, albeit it is interesting to see that for the largest elasticity of  $e = 0.5$ , the distributions lie to the right of the true elasticity of 0.5 suggesting the bunching approach, on average, overestimates the elasticity. The graphical evidence is supplemented in Table 1, which shows the RMSE for all specifications. Here, the data-driven procedure always outperforms the other two procedures in the sense that it delivers the smallest value. The increase in the RMSE is roughly between 10 and 25% when increasing the boundaries of the bunching window by two bins, which also holds for the subsequent increase by a further two bins.

Finally, we test three alterations to our simulation setup to provide evidence that our results hold for a variety of specifications. We test our data-driven procedure against two alternative bunching windows that are two and four bins larger on each side. Figure 3 plots the RMSE of these three bunching windows for each scenario, where the blue circle shows the result for our data-driven method, the red triangle shows the result for the two-bin larger window and the plus sign shows the RMSE for the case where we use a bunching window that is four bins larger on each side. Panel (a) of Fig. 3 provides the results of variations in the size of the optimisation frictions, which are implemented as a percentage of the kink point value. Having more optimisation frictions causes the bunching mass to be distributed along a larger stretch of the income distribution, resulting in a larger bunching window that should be harder to detect via eyeballing, as the distribution is smoother. Conversely, smaller optimisation frictions should result in a narrower bunching window. Panel

<sup>12</sup> As an example, if our procedure finds the bunching window to be from  $-3$  to  $+1$ , the comparison will be with  $-5$  to  $+3$  ( $-7$  to  $+5$ ). Please note that in an ideal setting, we would like to compare our method with the standard practice of “visual inspection”. As this is not feasible, we compare our method against two artificial windows that are two and four bins larger on each side (which follows the intuition that researchers typically choose windows that are too large).

(b) shows the results from varying the binwidth, and in Panel (c), we use a log-normal income distribution to hedge against any concerns that our results are driven by the choice of a triangular distribution in our baseline specifications. We set threshold values at  $k = 10,000$ ,  $k = 20,000$ ,  $k = 40,000$  and  $k = 60,000$ , which are located in various parts of the income distribution (see Fig. 7 of Appendix 2).

In line with our baseline results, Fig. 3 shows that, across all specifications, our data-driven method in determining the bunching window has the smallest RMSE in each comparison, i.e. the elasticity is more precisely estimated. Interestingly, our method performs very well when we move to the more realistic setup of a log-normal income distribution. Especially for the case where the kink is set at the convex part of the income distribution, our method delivers a much smaller RMSE compared to the other two bunching windows.

### 3 Institutional background

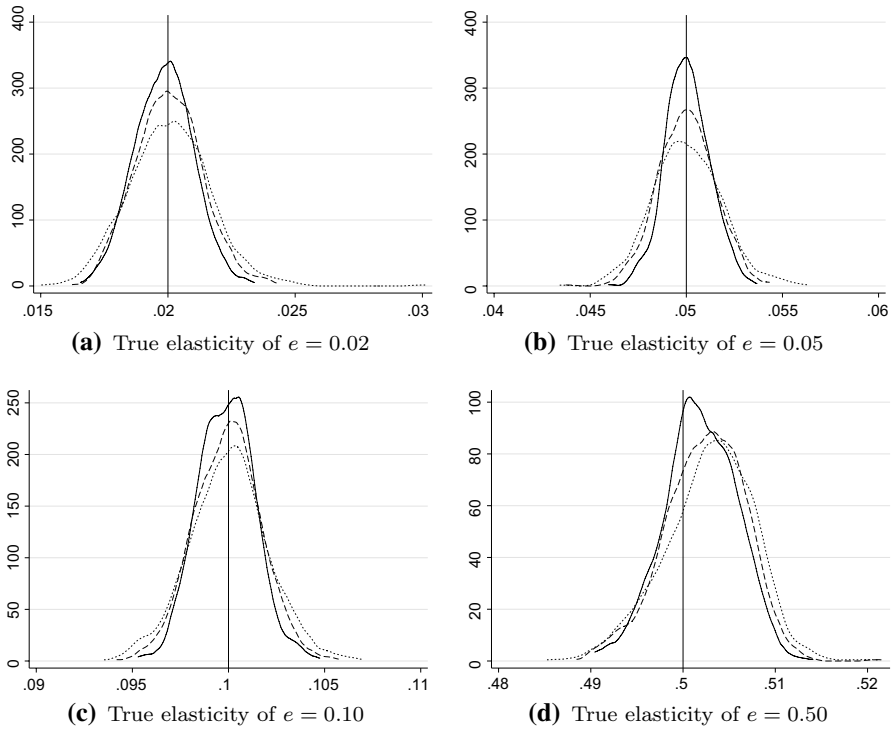
The Dutch tax system is almost fully individualised, and tax liabilities mainly depend on individual worldwide income. However, there are a few exceptions, two of which are relevant for our analysis. The first exception is that of means-tested subsidies, such as on health tax, child care and rent, which are all based on taxable household income. The second is that personal tax-favoured expenditures are transferable between partners, thus reducing taxable income. This last possibility is attractive under a progressive tax schedule.<sup>13</sup>

Since a major tax reform in 2001, income is treated in three different “boxes” dependent on its source, each with its own taxable income concept and tax schedule. In Box 1, income from profits, employment and home ownership is taxed. This includes wages, pensions and social transfers. Box 2 consists of income from substantial shareholding such as dividends and capital gains. Any other income from savings and investments is taxed in Box 3. Regarding our analysis, it is important to know that income losses in one box cannot be used to counterbalance taxable income in one of the others. Income in Box 1 is taxed at progressive rates that jump up at certain thresholds and thus create kinks in the tax schedule, whereas income in Box 2 and Box 3 is subject to a flat tax, that, in 2014, was 25% and 30%, respectively.<sup>14</sup> For our analysis, we exploit the kinks in the Box 1 tax schedule for identification.

Income in Box 1, minus personal deductions, is taxed at progressive tax rates. Tax rates in the first and second tax bracket also include a social security contribution of around 31% for old-age pensions and exceptional medical expenses. Figure 4

<sup>13</sup> From a labour supply perspective, a third exception is also relevant. A non-working spouse can transfer the lump-sum tax credit to his or her partner. The moment this spouse starts working, their income will be taxed starting at the marginal tax rate. This, however, is not the focus of our study.

<sup>14</sup> We are aware of the possibility of moving income between the boxes, which could be especially pronounced for self-employed individuals. For information on the importance of shifting between tax bases see Harju and Matikka (2016). Because of data limitations, we are unable to extend our analysis in that way.



**Fig. 2** Results Monte Carlo exercise. *Notes* The figures represent the results of our Monte Carlo simulations using 600 repetitions each. The solid line represents our data-driven procedure and the dashed (dotted) line the procedure where two (four) bins were added on each side

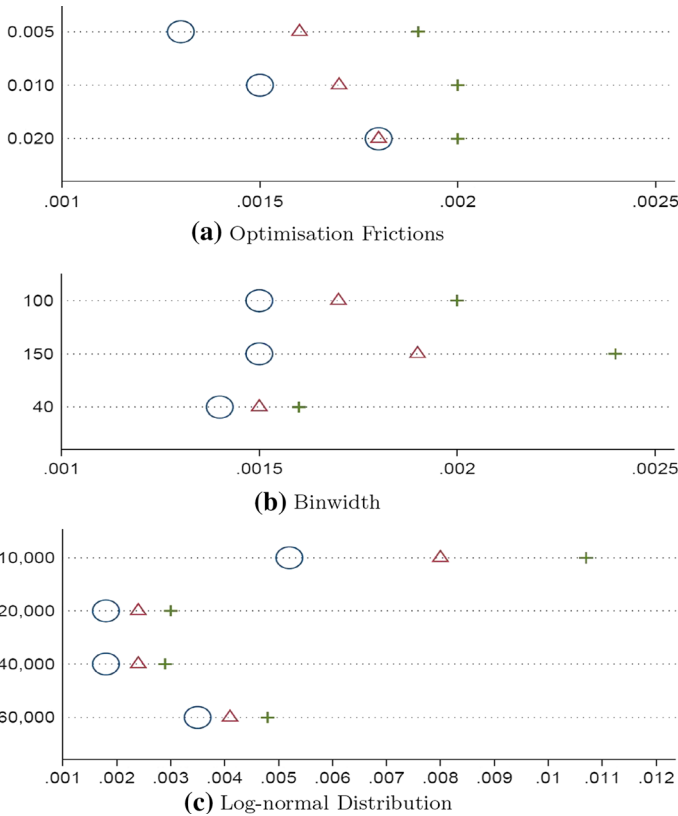
**Table 1** RMSE of simulated elasticities

	Elasticity			
	0.02	0.05	0.1	0.5
Data-driven	0.0011	0.0011	0.0015	0.0043
Data-driven+2 bins	0.0013	0.0015	0.0017	0.0049
Data-driven+4 bins	0.0016	0.0018	0.0020	0.0056

The table shows the RMSE for the three estimation procedures (data-driven, data-driven+2 bins and data-driven+4 bins) and four true elasticities

provides an snapshot of the Dutch tax schedule in 2014. The marginal tax rate is represented by the solid line. In 2014, there is an increase in the marginal tax rate of 8 percentage points at the first threshold and a large jump in the marginal tax rate, from 42 to 52%, in the last bracket.

While the upper two tax rates stayed constant over the whole sample period, there exists some variation over time for the lower two tax rates. Figure 8 in Appendix 2 graphically illustrates the development of each marginal tax rate over time, again adding the social security contributions for the lower two tax brackets. Due to the

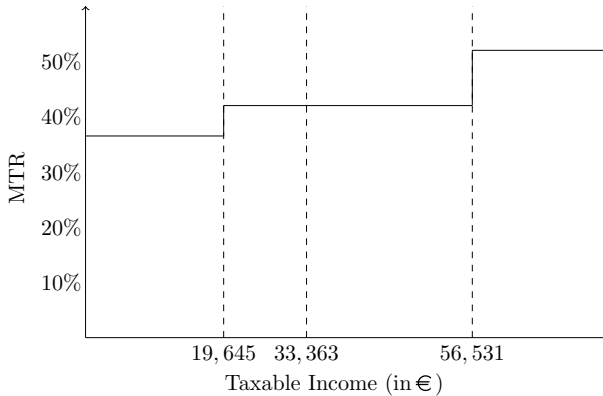


**Fig. 3** Monte Carlo simulation—RMSE robustness checks. *Notes* These three sub-graphs show the results of varying the size of optimisation frictions and binwidth in (a) and (b). In (c), we analyse the performance of our method for the case of a log-normal income distribution where we vary the location of the kink. The blue, hollow circles show the RMSE for our data-driven method, while the triangle and the plus show the case where a bunching window that is two and four bins larger on each side is used. All specifications have a true elasticity of  $e = 0.1$  (Color figure online)

stability of the two upper tax rates, the large jump of 10 percentage points at the highest threshold was existent in all years since 2001. It is also important to know that for the considered time period, the income thresholds were adjusted upwards to account for inflation and to avoid the phenomenon of “cold progression”<sup>15</sup> (see Table 4 in Appendix 2). However, given the specific values of the thresholds, which are never a multiple of one hundred, we are less concerned with round number bunching (Kleven and Waseem 2013).

To better understand how individuals adjust their taxable income, it is essential to know its exact definition. An overview of the computation of taxable income in the

<sup>15</sup> Cold progression (also known as “bracket creep”) describes the process by which individuals are shifted into higher tax brackets because tax thresholds are not adjusted for inflation. Stated differently, individuals have to pay more taxes although their real incomes have not increased.



**Fig. 4** Tax schedule of 2014, Box 1. *Notes* The figure shows marginal tax rates for the year 2014. At each threshold, denoted by the dashed lines, the marginal tax rate jumps up, except for the second threshold, where the tax rate and the social security contributions in the lower bracket equal the marginal tax rate in the higher bracket

Netherlands is given in Table 2. One important channel of adjusting taxable income is legal tax avoidance by utilising deductions (Chetty et al. 2011). In our setting, these deduction possibilities include alimonies paid, charitable giving, health expenditures or mortgage interest deductions.<sup>16</sup> In the Netherlands, the mortgage interest deduction is quite high and common among house owners. More importantly, all of these deductions can be shifted between partners.

Finally, important for any bunching analysis is the exact tax payment procedure which has an influence on the technical possibilities to avoid or evade taxes.<sup>17</sup> Three things are worth noting here. First, it should be emphasised that for people in paid employment, their employer withholds income tax from the income taxed under Box 1. This third-party reporting is important for the interpretation of the results as it makes systematic tax evasion—one way of adjusting taxable income—more difficult (Kleven et al. 2011). Second, an important distinction is single filing or joint filing of tax returns. Even though the Dutch tax system is rather individualised, married couples file their returns jointly. In addition, cohabiting couples are also allowed to jointly file their tax return, provided they have lived together for more than 6 months. Third, taxes can be filed digitally (computer-assisted) or on paper. The share of digital filers has increased dramatically from about 30% in 2003 to almost 95% in 2015. Digital filing of tax returns is not only helpful when deducting certain personal expenditures, but also facilitates the optimal shifting of income between partners. The exact threshold might become more salient and enables people to locate at the threshold.

<sup>16</sup> These are (at least in part) common in other countries like Great Britain or Germany.

<sup>17</sup> The tax thresholds in the Netherlands are known before the start of each fiscal year, as these are published together with the governmental budget which is presented each year, on the third Tuesday of September (Prinsjesdag).

**Table 2** Definition taxable income box 1

---

<i>Gross wage</i>	
– Pension fund and unemployment insurance contributions employee	
+ Health insurance contribution employer	
= <i>Taxable labour income</i>	
+ Income from housing	
+ Freelance earnings	
– Alimony/maintenance paid	
– Charity donations	
– Mortgage interest deductions	
– Health expenses deduction	
– Other personal deductibles	
= <i>Taxable income Box 1</i>	

---

This table shows the computation of Box 1 taxable income. Gross wage includes pension benefits and received social transfers

## 4 Data

The data used in this study are the income panel data (IPO) provided by Statistics Netherlands. This longitudinal data set covers the period from 2001 to 2014. It contains administrative data on all possible sources of income, on an individual level, as well as a very detailed account of possible deductions from the tax base. Most importantly for this study, Statistics Netherlands provides the information on relevant taxable income for Box 1 (see Table 2). The taxable income variable is obtained from the tax authority, representing the exact taxable income per individual. This circumvents the problem of measurement error, which is vital for bunching analyses. As our income measure includes all tax deductions, we do not have to rely on tax simulators that are used in other studies (Gruber and Saez 2002; Chetty et al. 2011).

The data set also includes demographic characteristics, which we exploit to study heterogeneity in the bunching behaviour of different socio-economic groups. We provide separate estimates for self-employed individuals, who theoretically would be more prone to bunching because of the lower costs and greater possibilities of adjusting their taxable income. Furthermore, we distinguish people according to gender and filing status. Our estimation sample is thereby restricted as follows: we exclude students as well as all people receiving governmental benefit payments. Because the tax is different for individuals aged 65 and over, we also exclude them from our estimation, as well as those below the age of 18. We omit the years 2001 and 2002 to avoid the inclusion of any after effects of the 2001 major Dutch tax reform. Furthermore, we only retain individuals with a positive reported taxable income. The pooled sample consists of  $N = 1,219,572$  individuals, which is roughly 1% of the Dutch population per year. The sample is evenly balanced with respect to gender (55% male) and married individuals (65%). Furthermore, the sample contains 14% self-employed individuals, including owners of small corporations, who would be in a position to decide on their own salary and are able to adjust it.

## 5 Results

### 5.1 Bunching evidence

Figure 5 gives a first glance at the bunching behaviour within the complete income distribution for the most recent sample year where the data are collapsed into income bins of 200 euros. The income thresholds of 2014 are indicated by vertical lines. Clear bunching behaviour can be seen at the first and third threshold, in line with the incentives set by the tax schedule.

We start out with the upper threshold, where the change in the marginal tax rate is largest with 10 percentage points (or 23.81%). Here, the incentive to bunch is most pronounced. Figure 6 reports the results for our pooled sample from 2003 to 2014 showing the number of observations per bin, relative to the threshold value. For this threshold, our proposed method delivers an asymmetric bunching window ranging from  $-483$  to  $+207$  euros.<sup>18</sup> In addition, the BIC criterion suggests a seventh-order polynomial counterfactual model for the upper threshold. In order to calculate an elasticity according to Eq. (4), a weighted average threshold value is used ( $k = 54,163$  euros). The weights are constructed by the number of taxpayers located exactly at the kink in each year. Standard errors are calculated using bootstrapping techniques.

We observe sharp bunching at the third threshold and estimate an excess mass of  $b = 1.67$ . The estimated excess mass translates into an ETI of 0.023, which is statistically significant at all usual significance levels. Quantitatively, a 10% decrease in the net-of-tax rate would induce a 0.23% reduction in taxable income. From an economic point of view, the tax response at this threshold is small, but in line with the recent findings of the bunching literature.

Following the simulation-based assessment of our data-driven procedure in comparison with larger bunching windows in Sect. 2.1, we first replicate the analysis depicted in Fig. 6 with a bunching window that is two (four) bins larger on each side. The results are shown in the second and third rows of Table 3. As apparent from the last two columns of this table, we are able to confirm the finding from the simulations that while all procedures deliver similar point estimates, i.e. there is no significant increase in the bias when using larger windows, the precision of the estimate drops considerably when utilising larger than optimal windows. The standard error of the estimate increases by around 25% (46%) when two (four) bins are attached to each side of the optimal bunching window. Thus, the efficiency of the estimator is severely hampered when a non-optimal bunching window is used in the estimation of the ETI.

In a second step, we also compare our data-driven procedure to determine the bunching window with the estimates of using two different, symmetric bunching windows, to further assess whether our procedure provides an improvement. The

---

<sup>18</sup> We use a 95% confidence interval for determining the bunching window throughout this study. We also tested a smaller confidence level, i.e. a one-standard deviation increase, which corresponds to a 68% confidence interval. The results are slightly larger but less precisely estimated.

bottom two rows of Table 3 display the results using a small bunching window from –345 euro to +345 euro and a large bunching window from –897 euro to +897 euro. Compared to Fig. 6, both the small and the large windows deliver smaller ETIs combined with larger standard errors, which supports the use of our data-driven methodology.

Finally, we estimate the excess mass for taxpayers and the ETI at the third threshold for all years separately. This eases any concerns about using a weighted average threshold in the pooled sample and hedges against possible bias stemming from observing individuals multiple times. The results are provided in Appendix 1 (Panel (a) of Table 5). One striking observation is that the bunching behaviour of individuals is increasing and becoming more precise over time. We ascribe this to learning effects, as taxpayers became more familiar with the tax system that fundamentally changed in 2001. Another explanation for this increase in the amount and precision of bunching could be the emergence of digital filing of tax returns, which could have made the threshold more salient to the general public.

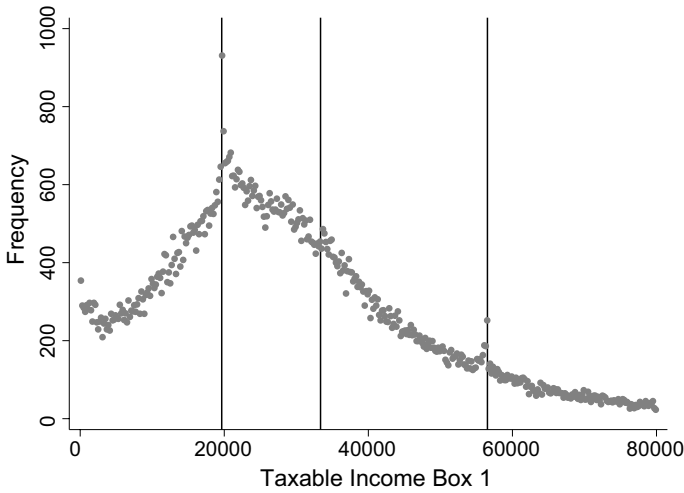
A case could be made for bunching at the other thresholds of the Dutch tax system as well, albeit that the change in the net-of-tax rate is much smaller and therefore we would expect less reaction. Figure 9 shows the results for the pooled sample for the first and second threshold, respectively. Surprisingly, we observe clear bunching behaviour of individuals at both thresholds.

At the first threshold, the income levels are quite low, which might suggest that individuals are more dependent on their income and should therefore show little real responses to a change in the marginal tax rate. However, the estimate for the ETI is about four times higher when compared to the third threshold ( $e_{th1} = 0.086$  vs.  $e_{th3} = 0.023$ ). Note that exact estimates for the ETI at both thresholds cannot be depicted, because we have changing tax differences over time, in addition to the changing threshold values. To calculate the ETI for these two thresholds, we used a weighted average tax change on top of the weighted average threshold. However, taking the average of the single-year ETI estimates as a sensitivity check delivers similar results: an elasticity of 0.085 at the first threshold. Contrary to our hypothesis, individuals with an income around the lower threshold seem to be more engaged in all kinds of tax-optimising behaviour in order to relocate at the threshold. In addition, the bunching behaviour is less precise and there is slightly more mass to the right of the threshold, suggesting that individuals are either less informed or less able to accurately adjust their income.

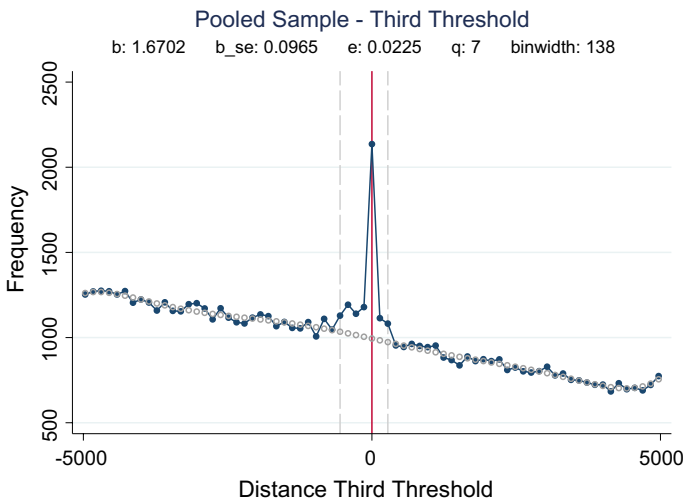
From an economic perspective, the second threshold might be of special interest for two reasons. First, the total change in the net-of-tax rate (tax rate plus social security contributions) is comparably small, with 3.35 percentage points as a maximum in 2003. The gain of manipulating taxable income thus may not be larger than potential adjustment costs, which would lead to less bunching. Second, the jump in marginal tax rates vanished in some years due to the adjustments of the tax rate in tax bracket 2.<sup>19</sup> Especially in these years, there is no incentive to bunch at the respective threshold. Despite these small incentives, the right graph of Fig. 9 clearly

<sup>19</sup> This phenomenon occurred in 2009, 2013 and 2014, respectively.





**Fig. 5** Income distribution in 2014. *Notes* This figure shows the sample distribution of income below 80,000 € in the Netherlands for 2014. The data are collapsed into 200-euro bins. The vertical lines represent the first, second and third threshold of the Dutch tax system, respectively



**Fig. 6** Bunching at the third threshold—pooled sample. *Notes* In this figure, bin counts are plotted relative to the third threshold for the pooled sample from 2003 to 2014. The bunching window is between -483 and +207 euros and the counterfactual model is a seventh-order polynomial (Color figure online)

shows that there is bunching behaviour at the second threshold. For the pooled sample, the estimated ETI amounts to 0.212 and is much higher compared to the other thresholds. This can be partly attributed to the small tax changes (frequently below 1 percentage point) that are used to derive the ETI. As expected, the result is largely driven by early periods of the sample, where the jump in the marginal

**Table 3** Estimation results bunching analyses—third threshold

	Window	bw	q	b	se	e
Data-driven	[- 4; 2]	138	7	1.6702	0.0965	0.0225
Data-driven+2	[- 6; 4]	138	7	1.6560	0.1209	0.0223
Data-driven+4	[- 8; 6]	138	7	1.6222	0.1408	0.0219
Symmetric small	[- 3; 3]	138	7	1.5128	0.0969	0.0204
Symmetric large	[- 7; 7]	138	7	1.6225	0.1539	0.0219

The table shows the estimates of the excess mass  $b$  and the respective elasticity of taxable income with respect to the statutory net-of-tax rate  $e$ , with *window* being the bunching window determined by our data-driven procedure. The binwidth is denoted by *bw*, and the order of polynomial is denoted by  $q$

tax rate was still noticeable. Figure 10 shows, for example, the results for the years 2008 and 2009, with 2009 being the first year where there was no jump at the second threshold. In 2008, we could still observe a small excess mass of 0.29, but in 2009, where the incentive to adjust taxable income vanished, we estimate a negative excess mass of merely  $-0.04$ , which is statistically insignificant.<sup>20</sup> For the year 2010, we are again able to identify a small excess mass of 0.18, which is significant at the 5% level. In accordance with the overestimation of small probabilities known from the behavioural economics literature (Kahneman and Tversky 1979), it might be the case that individuals overestimate benefits from small changes in the net-of-tax rate and subsequently adjust their taxable income even though the economic gain is minimal.

## 5.2 Subgroup analysis

In a last step, we split the sample according to gender, employment status and filing status to see whether there is heterogeneity in the bunching behaviour. The results are shown in Panel (b) of Table 5. The ETI is larger for women ( $e = 0.041$ ) than for men ( $e = 0.017$ ). Since there are considerably more men than women at the third threshold, this indicates that the ETI at the upper threshold in the pooled sample is predominantly driven by males. Given that there are roughly three times as many men than women around the third threshold, the ETI here in the pooled sample can also be described by the weighted sum of the elasticities of men and women, i.e.  $(3 \cdot e_{\text{men}} + e_{\text{women}})/4 = 0.023$ .

We then split the sample by employment status. Self-employed individuals have better possibilities to adjust their taxable income and are therefore more prone to bunching. The results confirm this hypothesis at the upper threshold. The ETI for self-employed is 0.042, which is about twice the size of the estimate in the full sample. In contrast to findings in many other studies, we also find a significant

<sup>20</sup> Note that because the tax change is zero in 2009, we cannot compute a value for the ETI and therefore argue via the excess mass.

elasticity for wage earners ( $e = 0.017$ ). One explanation for the significant bunching behaviour of employed individuals could be that of trade unions jointly setting wage levels for groups of individuals. Collective wage bargaining is very common in the Netherlands. The observed bunching for wage earners might also be an indication of collusion between employers and employees at the individual level and of contracts being specifically designed to achieve a taxable income at the threshold (Chetty et al. 2011). However, the degree of flexibility depends on the wage system. A majority of Dutch employees' wage payments are based on an industry-wide or company-wide wage schedule that resembles a staircase with fixed starting salaries and upper ceilings and fixed wage increments for each step in between (Deelen and Euwals 2014). On top of this general increase, employees and employers can decide on performance-related pay, which is at the heart of the individual wage differences.

Finally, we split the sample by filing status.<sup>21</sup> Married filers have the possibility to shift deductions between them, an option that is not available to single filers. A recent study by Doerrenberg et al. (2017) shows the importance of tax deductions for welfare analyses with the ETI. If individuals exploit the shifting of deductions as a main channel to adjust taxable income, we would expect to see no spike for single filers, while the estimates of the ETI for the joint filers should be close to the estimates from the pooled sample. The results shown at the bottom of Table 5 support this hypothesis. On the right, we can see no bunching behaviour of single filers, while for individuals that have the possibility to jointly file their tax returns, we clearly see bunching behaviour. The ETI is 0.026, which is similar to the 0.023 in the full sample.

Due to the high presence of mortgage interest deductions in the Netherlands, it is interesting to examine this special kind of deduction, which can only be claimed for one, usually the main mortgage. We analyse the anatomy of response of wage earners for 2011 for which year we have additional information on the shifting behaviour. Interestingly, almost 88% of all wage earners in the vicinity of the third threshold claim mortgage interest deductions. The sharp shifting between partners is visible in Fig. 11. The graph shows the share of total mortgage interest deduction claimed by one partner within a couple.<sup>22</sup> For high earners around the third threshold, the average share is around 80% of total mortgage interest deduction of the couple. The higher the income, the higher the share of the mortgage interest deduction claimed by the high-earning partner, which is in line with the common expectation that the tax advantage is higher for the highest earner. We would expect to see a jump at the third threshold, but we find a sharp dip in the share of mortgage interest deduction claimed by an individual at the upper threshold. This suggests that individuals strategically shift the mortgage interest deduction to their partners, as soon as they have located their taxable incomes at the threshold. Especially if both fiscal partners

<sup>21</sup> Married tax filers include both cohabiting unmarried couples who can choose to file tax returns together and married couples. In the Netherlands, married and cohabiting couples receive equal fiscal treatment.

<sup>22</sup> Unfortunately, as we have individual-level data, we cannot see whether it is the high-income partner who actually uses this type of deduction. However, the incentive to shift the mortgage interest rate is higher for the one with higher earnings.

earn more than the third threshold, this splitting of the mortgage interest deduction reduces the overall tax burden of the fiscal partners.

A second possible channel to reduce taxable income, other than shifting between partners, is that of a real response, for example in hours worked. Due to the structure of our data, identification of these types of responses in hours could only be done indirectly, for example, via hourly wages. We do not, unfortunately, observe actual working hours. As bunchers come from above the threshold and hourly wage can be assumed to increase with taxable income,<sup>23</sup> an individual that bunches should have a higher hourly wage than other individuals that obtain a similar taxable income. However, looking at data from 2006 to 2011 for which we have hourly wage data, we cannot detect any significant difference between bunchers and non-bunchers left or right of the bunching window in terms of hourly wages, suggesting that real responses do not play a significant role in adjusting taxable income.

### 5.3 Relation to the literature

Our results relate to the literature in several ways, although one has to keep in mind that cross-country comparisons of elasticities might be difficult due to different institutional features (Bastani and Selin 2014). In line with other studies that implement the bunching approach, we find small but precise estimates of the compensated ETI with respect to the net-of-tax rate at the top tax threshold of 0.023. Chetty et al. (2011) find an elasticity at the upper threshold below 0.02 for their full sample on Denmark, while Bastani and Selin (2014) find close-to-zero elasticities in Sweden at the top tax threshold. Evidence on the USA (Saez 2010) indicates an elasticity between 0.1 and 0.2, depending on the methodology, at the first threshold of the federal income tax schedule, and no response at other thresholds. For married couples, a large part of the response is driven by itemised deductions, whereas for singles income responses are most important. Saez (2010) notes that these income responses are harder to adjust than itemised deductions. We also find a stronger shifting response for married individuals than for singles. Still, income adjustments are easier in the Netherlands than in the USA because of the social acceptance and federal legitimization of part-time work. Employees in the Netherlands are arguably more free to choose their working hours than workers in other countries because of the existence of the Dutch Working Hours Act. In the USA, only 19% of the working population was working part-time in 2013, whereas in the Netherlands this figure was almost twice as high, at 36%. The significantly larger proportion of women bunching can also be explained by this. In the USA, 26% of the female workforce worked part-time, whereas in the Netherlands, this was 58% and these women would likely earn an income close to the first threshold.<sup>24</sup> This makes income adjustment through hours easier. In contrast to Saez (2010), we do find significant responses

<sup>23</sup> This can be justified for example by the higher skill level that high earners have compared to low earners.

<sup>24</sup> Shares are calculated from the OECD Statistics database, where the labour force is measured by national criteria.

at the higher tax brackets as well, which we attribute to the possibility to shift tax deductions between partners.

Earlier studies for the Netherlands find larger elasticities. Jongen and Stoel (2013) find an elasticity of around 0.1 for the short run and 0.2 for the medium run, with larger elasticities for women. The aforementioned study employs a panel approach and uses instrumental variable techniques to correct for endogenous taxes in line with Gruber and Saez (2002). In contrast to our study, they had to rely on a tax simulator to obtain marginal tax rates and determine taxable income. This can potentially cause measurement error, which could explain some of the deviation between the results. Another explanation would be that the bunching approach identifies a local elasticity as opposed to an average elasticity derived from the IV approach (Chetty 2012).

A recent study by Bettendorf et al. (2017) for owners of small corporations<sup>25</sup> finds elasticities between 0.06 and 0.11 for the upper threshold of the Dutch tax schedule, using bunching techniques. This is slightly larger than the elasticity of 0.04 that we identify for self-employed individuals, and could suggest that our results are partly driven by the DGA subgroup. Unfortunately, the limited number of owners of small corporations in our sample prevents us from running the estimation separately for this group.

In previous bunching studies, a distinction is made between real response and income shifting. In a study on the self-employed in Denmark, Le Maire and Schjerning (2013) show that about 50–70% of the bunching in taxable income is due to income shifting over time. In a similar study for business owners in Finland, Harju and Matikka (2016) attribute two-thirds of the ETI to income shifting between tax bases. However, we find that a large share of bunching is driven by tax deductions in combination with shifting them between partners and we find no indication for real responses.

## 6 Concluding remarks

We proposed a purely data-driven procedure to find an “optimal” bunching window. Monte Carlo simulations showed that our data-driven procedure outperforms larger bunching windows, especially in terms of precision (i.e. show the smallest standard error). This key result is, in addition, robust to various changes in key parameters, such as different binwidths, degrees of optimisation frictions and functional form assumptions.

We also employed a unique longitudinal data set containing exact declared taxable income for a representative sample of the Dutch population (IPO data from 2003 to 2014) to estimate the ETI. In our application to the Netherlands, we found an elasticity of 0.023 in the full sample at the upper threshold of the tax system, where the tax change is largest. Subgroup analysis revealed that self-employed individuals,

---

<sup>25</sup> These so-called DGAs (*Directeur-Groootaandeelhouder*) face a special tax scheme. In our study, this subgroup belongs to that of the self-employed.

women and married tax filers drive this result. The improvements, especially in efficiency, that were visible in the Monte Carlo simulation were also confirmed in the application: our proposed data-driven procedure outperformed both larger asymmetric bunching windows as well as two symmetric bunching windows in delivering a more precise estimate of the ETI.

In summary, our simple data-driven procedure does not only take away the researcher's discretion in setting the bunching window, but was also shown to improve the efficiency of the bunching estimator when compared to alternative techniques of determining the bunching window. Our modification thus fills a gap in the existing bunching literature.

**Acknowledgements** Open Access funding provided by Projekt DEAL. Nicole Bosch gratefully acknowledges funding from the Netherlands Organisation for Scientific Research (ORA-Project 464-11-091). We thank Abi Adams, Leon Bettendorf, Nadja Dwenger, Arjan Lejour, Nadine Riedel, Karsten Schweikert, Robert Stüber, Michael Trost and Mazhar Waseem for their invaluable comments, as well as the participants of the IIPF (2015) in Dublin, CPB seminar (2015) in The Hague, THE Workshop (2015) in Stuttgart, the Nordic Tax Workshop (2015) in Oslo, seminar at ZEW Mannheim (2016), EEA (2016) in Geneva, CBT Workshop (2016) in Oxford and EALE (2018) in Lyon for their helpful comments and discussion. We thank Statistics Netherlands for access to their microdata. Stata code is available from the authors upon request.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix 1: Additional material on the bunching (window) method

### Bunching method

The bunching theory underlying the empirical applications starts with the simple neoclassical model where individuals have well-behaved preferences over consumption and leisure and where discontinuities in marginal tax rates create kinks or notches in their budget sets. Although it was initially developed in the field of taxation (Saez 2010), it has been applied to many other research question in the area of social insurance, energy economics, education and sports economics (see Kleven (2016) for an overview). Since our application uses changes in personal income tax rates, the theory is explained in terms of earnings responses to taxation, but our methodological contribution can be applied to any other setting as well.

To test the prediction from microeconomic theory and to quantify the welfare losses due to income taxation, we follow the literature and identify the compensated ETI in the spirit of Feldstein (1995), which is a summary statistic for all kinds of behavioural responses (real responses, tax avoidance and tax evasion). The ETI is

hereby defined as the percentage change in taxable income  $z$  due to an increase in the net-of-tax rate  $(1 - \tau)$  of one percentage:

$$e(z) = \frac{dz}{z} \bigg/ \frac{d(1 - \tau)}{(1 - \tau)}. \tag{3}$$

The introduction of a kink in the budget set of individuals induces bunching behaviour within a certain income range, provided that preferences are convex and smoothly distributed among the population. This will lead to a spike in the income density exactly at the kink, but due to the inability of individuals to control or predict their incomes or imperfect information about the exact threshold value, a bunching window around the kink is observed more often in reality (Saez 2010).

Comparing the income density with a counterfactual scenario without a kink, one can show that the excess mass of taxpayers around the threshold can be exploited to determine the elasticity  $e(z)$  (Saez 2010). The compensated ETI, identified locally at the threshold  $k$ , is then given by

$$e(k) = \frac{b}{k \cdot \log\left(\frac{1-\tau_1}{1-\tau_2}\right)}, \tag{4}$$

where the net-of-tax rate changes by  $\log\left(\frac{1-\tau_1}{1-\tau_2}\right)$  percentage.<sup>26</sup> The relative excess mass of taxpayers at the threshold  $k$  is given by  $b$ , which is the only parameter that needs to be estimated. A value of  $b = 2$ , for example, corresponds to twice more individuals being at the threshold than would have been the case in the absence of any tax change. In short, the main idea behind the bunching approach is that the observed amount of excess bunching is informative about the (compensated) income elasticity.

**Formal derivation bunching window**

The bunching window is formally derived as follows: let  $x_- \in \{-X, (-X + 1), \dots, 0\}$  and  $x_+ \in \{0, 1, \dots, X\}$  be the respective lower and upper bound of the excluded region, where  $X$  represents the midpoint of an income bin. Furthermore, define  $l(x_-, x_+)$  as the lower bound of the bunching window and  $u(x_-, x_+)$  as the upper bound, given the excluded region from  $[x_-, x_+]$ .<sup>27</sup> For every tuple  $(x_-, x_+)$ , run a local regression of polynomial order  $q$ :

<sup>26</sup> It is identified if and only if the derivative of the counterfactual density function  $h_0(z)$  with respect to  $z$  is continuous in  $z \forall z$ .

<sup>27</sup> Please note again that the excluded region runs through all possible combinations starting from  $(0, 0)$  such that the excluded region is not always symmetric around the threshold. To give an example, using an excluded region going from  $-15$  to  $15$  would thus result in  $(15 + 1)^2 = 256$  regressions. Adding  $-16$  and  $+16$  would increase the number of regressions by 33.

$$\tilde{N}_j^{BW} = \sum_{i=0}^q \beta_i X_j^i + \varepsilon_j \quad \forall j \notin [x_-, x_+]. \tag{5}$$

Then, predict the counterfactual values  $\hat{N}_j^{BW}$ :

$$\hat{N}_j^{BW} = \sum_{i=0}^q \hat{\beta}_i X_j^i \quad \forall j. \tag{6}$$

As a next step, calculate the upper value of the confidence interval  $CI_j^+$  for a given  $t$ -value using standard procedures. To determine whether there are more individuals than predicted in an income bin  $j$ , subtract the  $CI_j^+$  from the observed number of taxpayers for each  $j$ :

$$E_j = N_j - CI_j^+. \tag{7}$$

A positive  $E_j$  means that the number of individuals in income bin  $j$  exceeds the predicted number of individuals, as estimated by the polynomial regression. Put differently, if all  $E_j$  are negative, no bunching is present in the respective sample. Otherwise, the lower bound of the bunching window is given by:

$$l(x_-, x_+) = j_l^* + 1, \quad \text{where } j_l^* = \max\{j \in \mathbb{Z}_- : E_j < 0\}, \tag{8}$$

which is the smallest contiguous income bin  $j$  that still satisfies the condition  $E_j > 0$ . Similarly, the upper bound is given by:

$$u(x_-, x_+) = j_u^* - 1, \quad \text{where } j_u^* = \min\{j \in \mathbb{Z}_+ : E_j < 0\}, \tag{9}$$

which is the largest contiguous income bin  $j$  that still satisfies the condition  $E_j > 0$ . Following this procedure, two distributions of lower and upper bounds of the bunching window are obtained. Ideally, every excluded region that is tested will deliver the same optimal bunching window resulting in a deterministic distribution. But especially when considering excluded regions that are within the true bunching window, the confidence interval around the prediction could deliver a wrong bunching window. Because the mean is affected by such outliers, we advocate to use the mode or median<sup>28</sup> to determine the optimal bunching window going from  $l^*$  to  $u^*$ .

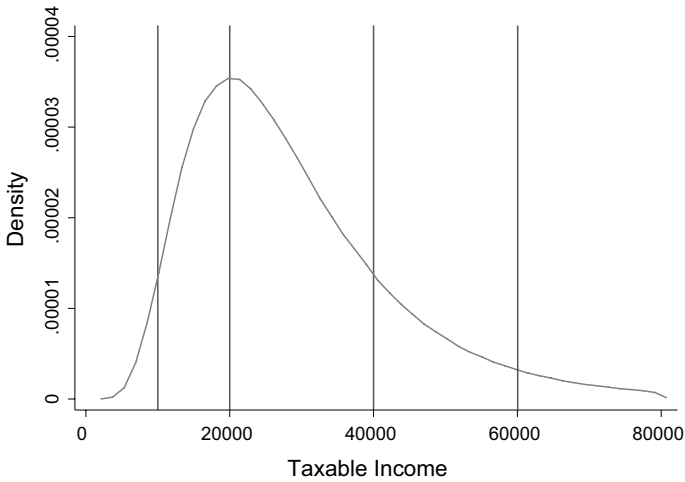
## Appendix 2: Additional graphs and tables

See Figs. 7, 8, 9, 10 and 11 and Tables 4 and 5.

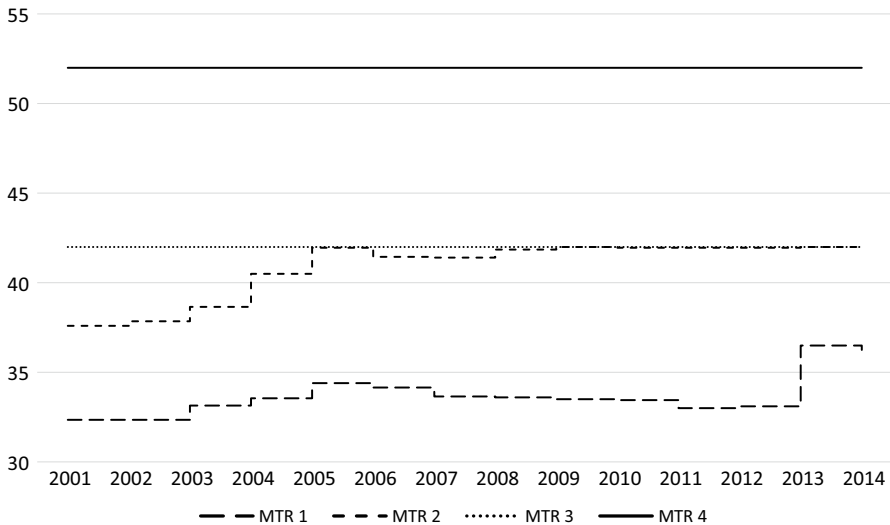
---

<sup>28</sup> In our simulation, all three procedures (mean, mode, median) finally delivered the same optimal bunching window for a given true elasticity.





**Fig. 7** Sample income distribution for Monte Carlo simulation. *Notes* The figure shows the income distribution used in the Monte Carlo simulations. The analysed thresholds at 10,000, 20,000, 40,000 and 60,000 euros are depicted by the vertical lines. The distribution is cut-off at 80,000 euros, which is roughly the 99th percentile of the simulated income distribution

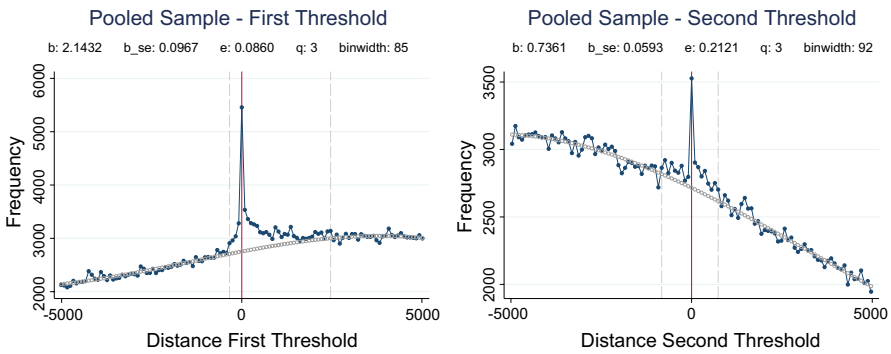


**Fig. 8** Development of marginal tax rates. *Notes* The figure depicts the development of the marginal tax rates in the Netherlands from 2001 to 2014

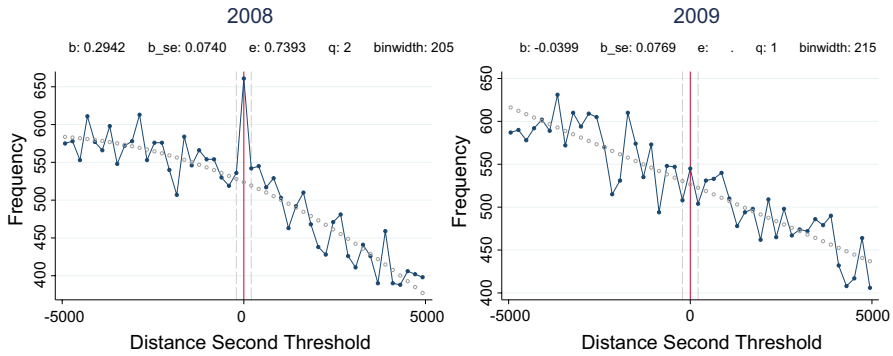
**Table 4** Development of thresholds

	First threshold	Second threshold	Third threshold
2001	14,870	27,009	46,309
2002	15,331	27,847	47,746
2003	15,883	28,850	49,464
2004	16,265	29,543	50,652
2005	16,893	30,357	51,762
2006	17,046	30,631	52,228
2007	17,319	31,122	53,064
2008	17,579	31,589	53,860
2009	17,878	32,127	54,776
2010	18,218	32,738	54,367
2011	18,628	33,436	55,694
2012	18,945	33,863	56,491
2013	19,645	33,363	55,991
2014	19,645	33,363	56,531

The table reports the development of the thresholds of the Dutch tax system from 2001 to 2014



**Fig. 9** Bunching at the first and second thresholds—pooled sample. *Notes* The figures show bunching at the first and second thresholds for the pooled sample from 2003 to 2014. The bunching window is between  $- 297.5$  and  $+ 2422.5$  euros for the first threshold and between  $- 782$  and  $+ 690$  euros for the second threshold. The counterfactual model is a third-order polynomial in both cases (Color figure online)

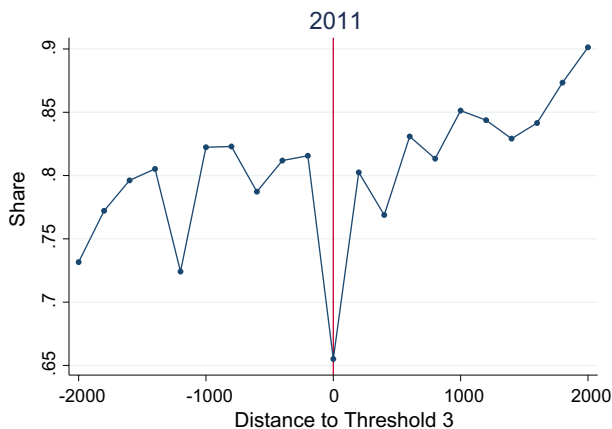


**Fig. 10** Bunching at the second threshold—2008 and 2009. *Notes* The figures show (non-)bunching behaviour at the second threshold of the Dutch tax system for the years 2008 and 2009. Because the tax change was zero in 2009 and it is part of the denominator in the elasticity formula, no value for the ETI can be estimated in 2009 (Color figure online)

**Table 5** Additional bunching results third threshold

	Window	bw	q	b	se	e
<i>Panel (a): Single years</i>						
2003	[− 1; 1]	351	1	0.2033	0.1345	0.0076
2004	[− 1; 1]	353	1	0.5609	0.1258	0.0207
2005	[− 1; 1]	350	3	0.5658	0.1650	0.0202
2006	[− 1; 1]	347	1	0.6609	0.1233	0.0232
2007	[− 1; 1]	329	1	0.5801	0.0864	0.0190
2008	[− 2; 1]	311	1	0.8384	0.1129	0.0256
2009	[− 1; 1]	303	1	0.9068	0.0867	0.0265
2010	[− 1; 1]	297	1	0.5641	0.1131	0.0163
2011	[− 2; 1]	301	2	0.9132	0.1156	0.0261
2012	[− 1; 1]	289	1	0.9335	0.1143	0.0252
2013	[− 1; 1]	316	1	0.7471	0.1037	0.0223
2014	[− 2; 1]	290	1	1.2876	0.1234	0.0349
<i>Panel (b): subgroups</i>						
Men	[− 3; 1]	152	1	1.1391	0.0769	0.0170
Women	[− 2; 1]	210	5	1.9902	0.1406	0.0407
Wage earners	[− 3; 1]	150	7	1.2066	0.0738	0.0176
Self-employed	[− 4; 7]	227	2	1.9069	0.1673	0.0423
Single filers	[− 1; 1]	332	2	0.1084	0.1005	0.0035
Joint filers	[− 4; 2]	142	1	1.9021	0.0788	0.0263

The table shows the estimates of the excess mass  $b$  and the respective elasticity of corporate taxable income with respect to the statutory net-of-tax rate  $e$ , with  $window$  being the bunching window determined by our data-driven procedure. The binwidth is denoted by  $bw$  and the order of polynomial is denoted by  $q$ . Panel a) shows single-year estimates; panel b) shows subgroup analysis



**Fig. 11** Share mortgage interest deductions. *Notes* The figure shows the share of mortgage interest deductions within couples around the third threshold for the year 2011. The binsize is 200 euro (Color figure online)

## References

- Adam, S., Browne, J., Phillips, D., & Roantree, B. (2015). Adjustment costs and labour supply: Evidence from bunching at tax thresholds in the UK. *mimeo*.
- Bastani, S., & Selin, H. (2014). Bunching and non-bunching at kink points of the Swedish tax schedule. *Journal of Public Economics*, *109*, 36–49.
- Bettendorf, L., Lejour, A., & van't Riet, M. (2017). Tax bunching by owners of small corporations. *De Economist*, *165*, 411–438.
- Blomquist, S., & Newey, W. (2017). *The bunching estimator cannot identify the taxable income elasticity*. Working Paper Series, No. 24136, National Bureau of Economic Research (NBER).
- Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, *80*, 969–1018.
- Chetty, R., Friedman, J. N., Olsen, T., & Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records. *The Quarterly Journal of Economics*, *126*, 749–804.
- Deelen, A., & Euwals, R. (2014). Do wages continue increasing at older ages? Evidence on the wage cushion in the Netherlands. *De Economist*, *162*, 433–460.
- Devereux, M. P., Liu, L., & Loretz, S. (2014). The elasticity of corporate taxable income: New evidence from UK tax records. *American Economic Journal: Economic Policy*, *6*, 19–53.
- Doerrenberg, P., Peichl, A., & Siegloch, S. (2017). The elasticity of taxable income in the presence of deduction possibilities. *Journal of Public Economics*, *151*, 41–55.
- 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*, 551–572.
- Freedman, D., & Diaconis, P. (1981). On the histogram as a density estimator: L2 theory. *Probability Theory and Related Fields*, *57*, 453–476.
- Gruber, J., & Saez, E. (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics*, *84*, 1–32.
- Harju, J., & Matikka, T. (2016). The elasticity of taxable income and income-shifting: What is “real” and what is not? *International Tax and Public Finance*, *23*, 640–669.
- Jongen, E. L. W., & Stoel, M. (2013). *Estimating the elasticity of taxable labour income in the Netherlands* CPB Background Document.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*, 263–291.
- Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, *8*, 435–464.

- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, *79*, 651–692.
- Kleven, H. J., & Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. *The Quarterly Journal of Economics*, *128*, 669–723.
- Le Maire, D., & Schjerning, B. (2013). Tax bunching, income shifting and self-employment. *Journal of Public Economics*, *107*, 1–18.
- Lediga, C., Riedel, N., & Strohmaier, K. (2018). The elasticity of corporate taxable income—Evidence from South Africa. *Economics Letters*, *175*, 43–46.
- Neisser, C. (2017). *The elasticity of taxable income: A meta-regression analysis*. IZA Discussion Papers, No. 11958, Institute of Labor Economics (IZA), Bonn.
- Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, *2*, 180–212.
- Saez, E., Slemrod, J., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, *50*, 3–50.
- Slemrod, J. (1998). Methodological issues in measuring and interpreting taxable income elasticities. *National Tax Journal*, *51*, 773–788.
- Weber, C. (2016). Does the earned income tax credit reduce saving by low-income households? *National Tax Journal*, *69*, 41–76.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.