Online Appendix

1   Bunching

A classical model predicts bunching at tax kinks when the budget set is convex, because individuals above the tax kink wish to decrease their income as the tax rate above the kink rises given their preferences (e.g. labor-leisure preferences in the context of earned income and saving-consumption preferences in the context of dividend and interest income). However, once they reach the tax kink, they have no incentive to decrease their income further because the tax rate below the kink did not change and they were willing to earn all income at this tax rate before the tax rate increased above the kink. If these individuals were able to perfectly adjust their AGI, they would all be located at a point mass at the kink; however, adjustment costs and income volatility generate imperfect bunching around the tax kink. Saez (2010) shows that bunching in the region of a tax kink can be used to calculate an elasticity of AGI with respect to the marginal net-of tax rate using this formula:

$$\hat{\varepsilon} \simeq \frac{\hat{B}/\hat{h}_0(z^*)}{z^* ln\left(\frac{1-\tau_l}{1-\tau_h}\right)}, \tag{1}$$

where z^* is the location of the tax kink, τ_l and τ_h are the marginal tax rates on either side of z^*, such that $\tau_l < \tau_h$. Let there be a region $[z^* - \delta_b, z^* + \delta_b]$ in which all imperfect bunching occurs. Then, $\hat{B}$ is the estimated excess mass and $\hat{h}_0(z^*)$ is the estimated counterfactual density of taxpayers in this region; that is, the density of taxpayers if there was no tax kink. These parameters are depicted graphically in Figure 1. Saez (2010) implements this method to estimate EITC elasticities at each kink; however, his analysis ignored that individuals' location relative to the second kink is usually determined by AGI, not earned income, as stressed in this paper.

To estimate $\hat{B}$ and $\hat{h}_0(z^*)$, I first estimate the smoothed density $\hat{h}$ using a local linear

regression within \$12,000 of the tax kink.[1] I use the smoothed density estimates to construct the elasticities, because this method provides substantial efficiency gains, relative to constructing the estimates using unsmoothed estimates, which is the method used by Saez (2010) and Kleven and Waseem (2011). I use an automatic bin and bandwidth selection criteria[2] to select the bin and bandwidth size within $[z^* - \delta_b, z^* + \delta_b]$ and apply this bin and bandwidth to the entire income range.[3] As long as the bin size chosen is small relative to the selected bandwidth, the exact choice of bin size is usually not important. However, Saez (2010) effectively over-smooths the data by using a bin size (\$500) that is an order of magnitude larger than the bin size used here for his smoothed density figures. Over-smoothing the data induces a downward bias in my elasticity estimates; for example, using the baseline bandwidth reported in this paper for self-employed individuals for all years, but replacing the bin size with \$500, biases the estimates downward by more than 10 percent, with a minimal decrease in variance.

Let $\hat{h}^*_-$ be the estimated mean density in the region $[z^* - \delta_b - \delta_c, z^* - \delta_b]$, $\hat{h}^*$ be the estimated mean density in the region $[z^* - \delta_b, z^* + \delta_b]$, and $\hat{h}^*_+$ be the estimated mean density in the region $[z^* + \delta_b, z^* + \delta_b + \delta_c]$, where δ_c is the width of the counterfactual region on either side of the imperfect bunching region. Let $\hat{H}^*_- = \delta_c \hat{h}^*_-$, $\hat{H}^* = 2\delta_b \hat{h}^*$, $\hat{H}^*_+ = \delta_c \hat{h}^*_+$, where these variables denote the cumulative density in their respective regions. Ideally, $\hat{h}^*_+$ would be the counterfactual density, absent a tax kink, in the region $[z^* + \delta_b, z^* + \delta_b + \delta_c]$. However, assuming τ_h is the tax rate above the tax kink, it overestimates the counterfactual if individuals above z^* decrease their income by a certain percentage in response to the higher tax rate (as opposed to decreasing their income by a fixed amount); this is consistent with what our models predict and the parameter our estimation strategies are designed to

---

[1] The figures are trimmed from -\$9,000-\$9,000 to exclude the edges which are noisy because the estimates are smoothed over a small number of observation at the edges.

[2] The bandwidth is based on the rule-of-thumb approach given in Fan and Gijbels (1996). This is the method used in McCrary (2008), whose procedure is the same as mine, except I do not allow for a break at z^* because imperfect bunching should take place on *both* sides of the tax kink.

[3] A bin size and bandwidth chosen over a larger region would not efficiently measure the density in the imperfect bunching region.

uncover. It overestimates the counterfactual above the tax kink because the counterfactual distribution gets compressed as each individual decreases their income by a percent of their total income. Therefore, the elasticity estimates of a particular income type with respect to the marginal net-of-tax rate $(1 - \tau)$ presented in this section can be interpreted as lower bounds on the truth.[4]

The counterfactual density is the mean of the densities above and below the region of imperfect bunching: $\hat{h}_0(z^*) = \frac{1}{2}(\hat{h}_-^* + \hat{h}_+^*)$. This method assumes that the densities on either side of the region of imperfect bunching are a good approximation for the counterfactual density in the imperfect bunching region.[5] The excess mass is constructed by subtracting the counterfactual mass from the total mass in the bunching region: $\hat{B} = \hat{H}^* - \frac{\delta_b}{\delta_c}(\hat{H}_-^* + \hat{H}_+^*)$.

The other important consideration in constructing these estimates is the choice of the size of the imperfect bunching region, $2\delta_b$, and the size of the counterfactual region, δ_c. I set $\delta_b = \delta_c = \$1,000$ for the baseline estimates, which are the same as those used by Saez (2010) to analyze the second tax kink. The choices for δ_b and δ_c are based on the following considerations. First, the gap between the first and second kink for two-child families since 1996 is $\$3,470$.[6] Therefore, it must be that $2\delta_b + \delta_c \leq 3,470$, so that the counterfactual region for each kink does not include part of the bunching region for the other tax kink. If δ_b is smaller than the imperfect bunching region, the estimates will be biased downwards because part of the bunching will be excluded from $\hat{H}^*$ and will instead be included in $\hat{H}_+^*$. Assuming the counterfactual density is flat, choosing a δ_b that is larger than the imperfect bunching region does not bias the estimates, but will otherwise. In practice, choosing a large

---

[4]Previous literature has either ignored this issue (Saez, 2010; Kleven and Waseem, 2011) or assumed that the density is biased downwards because the bunching individuals came from the region above the tax kink (Chetty et al., 2011). The latter assumption is incorrect as long as individuals further above the tax kink are responding in the same way as those near the tax kink and the counterfactual region is not near the next tax kink (here there is a decline because these individuals are no longer being replaced by individuals further up in the distribution).

[5]This is the most common way of constructing the counterfactual density in the literature (Saez, 2010; Kleven and Waseem, 2011). The other approach is to estimates a global polynomial to construct the counterfactual (Chetty et al., 2011). This approach is not ideal here because there are two kinks that are relatively close together and the kinks are not defined by the same type of income.

[6]The gap increases for married families in 2002 by $\$1,000$ and by another $\$1,000$ in 2005.

δ_b induces a bias in the estimates, because the extent to which $\hat{h}^*_-$ and $\hat{h}^*_+$ provide an accurate estimate of $h_0(z^*)$ usually declines further away from z^*. There is not much that can be done in the way of sensitivity analysis over the whole sample period given the constraint that the first kink imposes. I pursue this issue below, when I look exclusively at the years 1988-1993, when the first kink was about twice as far away as it was in 1994-2006.

I examine self-employed individuals and wage-earners separately because they arguably differ in their capacity to respond to tax kinks. An individual is defined as self-employed if they have any self-employment income. One way in which individuals can bunch at the tax kink is to not report income above the kink—tax evasion. It is difficult to evade taxes on wage and salary income without being caught, as this income is generally subject to withholding and information reporting.[7] Self-employment income faces neither withholding nor information reporting.

Tax avoidance is another way in which individuals can respond. There are few tax avoidance possibilities on earnings for wages because wages face withholding and information reporting requirements. Still, there is some flexibility if workers are able to substitute towards non-monetary forms of compensation. Self-employed individuals have a greater opportunity for tax avoidance. These individuals are allowed to deduct their expenses from their gross income, which allows them to have an extra expense, say, which would place them at the tax kink without engaging in anything illegal.

A labor supply response could also move these individuals towards the tax kink. Self-employed individuals face lower adjustment costs, on average, associated with altering their labor supply, so it is more feasible for them to make minor changes.

Each of these three possibilities—evasion, avoidance, and a labor supply response—all

---

[7]Note that there is a minimum requirement on withholding for wage-earning individuals. It is $221 for single individuals and $667 for married individuals in 2006 dollars. This should, in general, not have an effect on the estimates, unless individuals hold two jobs, one of which is below the threshold and they claim to only have one job on the Form W-4 for their primary job (Form W-4 is completed by individuals so that their employer knows how much to withhold). If individuals decide to report the income from their second job if they face the lower tax rate, but not if they face the higher rate, a small part of the response by these individuals could be tax evasion.

suggest that the income adjustments will be larger and more precise for the self-employed compared to wage-earners. The top panels of Figure 2 display the bunching at the second EITC kink for self-employed and wage-earning individuals for all years. The densities are population-weighted and normalized so that they integrate to one. The variables $\hat{z}^*$, $\hat{\tau}_l$, and $\hat{\tau}_h$ are calculated as the population weighted averages for each sample. The solid yellow line below the density is the estimated counterfactual density of individuals in the bunching region. The bunching depicted in the figures is more sharply defined for the self-employed. The corresponding elasticities are given in Table 1 Columns (1) and (2). Standard errors for the elasticities were obtained by nonparametric bootstrap (1000 replications) clustered by state. The elasticity estimate of AGI with respect to the marginal net-of-tax rate for self-employed individuals is 0.063 and is significant at the one percent level, while the estimate for wage-earners is 0.010 and insignificant.

Table 1 Columns (3) - (5) and the lower panels of Figure 2 examine years 1988-1993 for self-employed and wage-earners, respectively. The key advantage of examining the early years of the tax credit is that the first kink was about twice as far away for filers with two or more dependents, so a broader range of δ_b and δ_c are feasible.[8] I fix $\delta_c = \$1,000$. For the self-employed, I report estimates for $\delta_b = \$2,000$ and $\delta_b = \$1,000$ in Columns (3) and (4); the counterfactual density is drawn in Figure 2 for the case of $\delta_b = \$2,000$. The estimates are similar for both and visually $\delta_b = \$2,000$ appears correct, but there are efficiency gains when using $\delta_b = \$1,000$. The elasticity estimate for $\delta_b = \$1,000$ is 0.093 and is marginally insignificant at the 10 percent level. One could even make the case that δ_b should be $3,000. In this case, the estimate roughly doubles and this rise is driven by the fall in density between -$4,000 and -$3,000, which is now part of the counterfactual density. For wage-earners, the

---

[8]Also, in the later years, the second kink interacted with several other tax schedule features. The filing threshold for married individuals was within $1,000 of the second kink in 1994-2006. However, it is unlikely that this had a substantial effect on filing, as it was best for individuals on both sides of the tax kink to file to claim the EITC. In 2002, the Saver's Credit was introduced and the first notch for single filers is near the second EITC kink. Ramnath (2010) documents bunching at these kinks, but the number of individuals applying for the Saver's Credit was small, so it is unlikely that this would create a substantial bias in these estimates. If either of these programs were playing a significant role in the size of the overall estimates, the estimates would be biased upwards.

plotted density strongly supports $\delta_b = \$3,000$.

The fact that the bunching is more spread out for wage-earning individuals is consistent with the fact that they have less flexibility in adjusting their earned income. The elasticity estimate for wage-earners is given in Column (5). It is 0.104 and is statistically significant at the five percent level. The standard errors on the wage-earner estimates are more precise because there are many more wage-earning than self-employed individuals. The wage-earner and self-employed elasticities are about the same (once the difference in the precision of bunching is taken into account) for these years. There are several possible explanations for this finding. First, when self-employed individuals engage in tax evasion, they may move all the way to the first tax kink (in the region between the first and second kink, they receive no additional EITC benefit, but do have to pay additional payroll tax on their self-employment earnings). Alternatively, while the self-employed may find it easier to alter their earned income, wage-earners may make up the difference by adjusting their unearned income more. The estimates in this section can be interpreted as a lower bound relative to what would be found further in the phase-out region, where the amount of awareness needed to respond precisely is lower.

References

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, 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.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and its Applications*. Chapman & Hall, New York.

Kleven, H. J. and Waseem, M. (2011). Tax notches in Pakistan: Tax evasion, real responses, and income shifting. Working paper.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142:698–714.

Ramnath, S. P. (2010). Taxpayers' response to notches: Evidence from the Saver's Credit. Working Paper.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212.

Figure 1: Estimating Elasticity from Bunching Around a Tax Kink
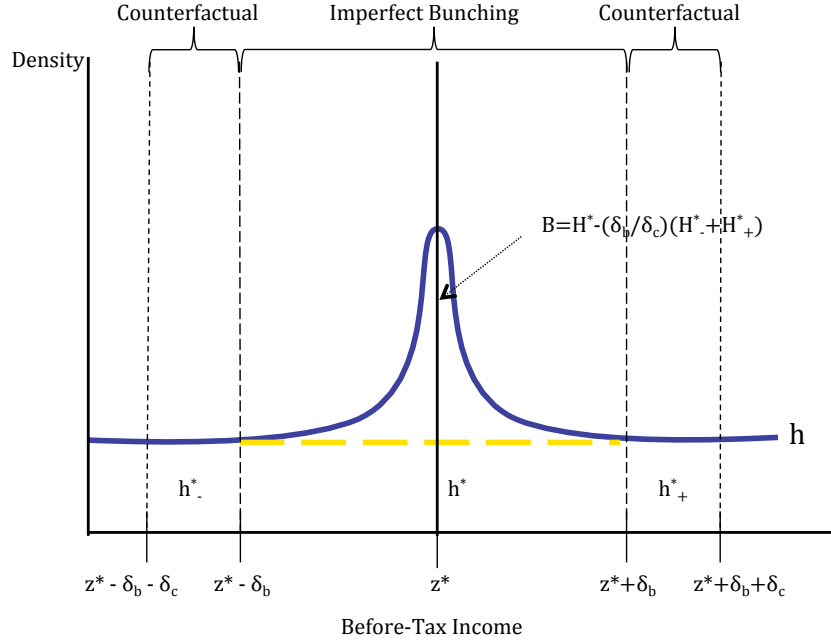
This figure illustrates how bunching at a tax kink can be used to identify the response to the change in marginal tax rate at the tax kink. z^* is the tax kink, $2\delta_b$ is the width of the imperfect bunching region, δ_c is the width of the counterfactual region. h^*_- and h^*_+ are the densities in the counterfactual regions, h^* is the density in the bunching region, each of which is depicted by the blue line in its respective region. $H^* = 2\delta_b h^*$ is the cumulative density in the imperfect bunching region. $H^*_- = \delta_c h^*_-$ and $H^*_+ = \delta_c h^*_+$ are the cumulative densities in the counterfactual region. B is the amount of imperfect bunching in the bunching region, which is given by the area between the solid blue and dashed yellow lines.

Figure 2: Estimated Density Around the Second EITC Kink for Self-Employed Individuals
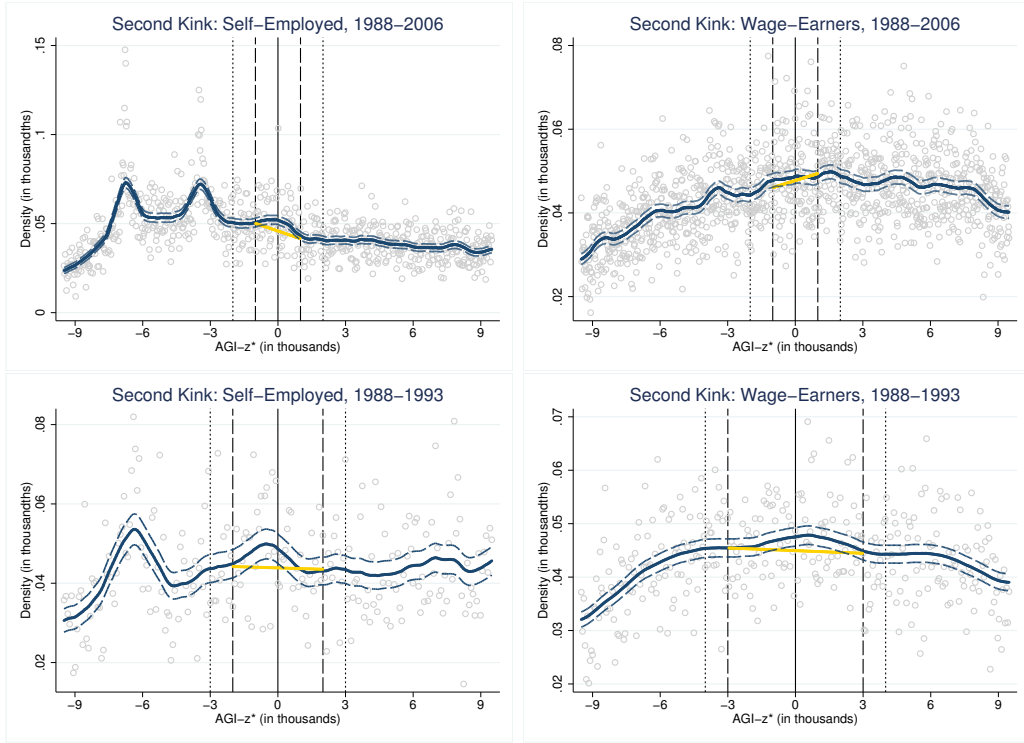
All adjusted gross income (AGI) values are in 2006 dollars, renormalized so that the kink (z^*) is at zero for all years. The estimates are population weighted. The density estimates are normalized to integrate to one. 95 percent confidence intervals are given by the dashed lines. The scatter-plot provides the unsmoothed density estimates. The region inside the long-dash vertical lines is used to calculate the bunching at the tax kink, and the region between the short-dash and long-dash lines is used to calculate the counterfactual density. The implied counterfactual is given by the solid yellow line. The bin size and bandwidth are 27 and 797 for the top left panel and are 83 and 1,469 for the bottom left panel. The bin size and bandwidth are 17 and 745 for the top right panel and are 55 and 2,310 for the bottom right panel.

Table 1: Bunching Elasticities at Second EITC Kink

	(1)	(2)	(3)	(4)	(5)
Elasticity	0.063***	0.010	0.104	0.093	0.104*
	(0.026)	(0.016)	(0.080)	(0.067)	(0.055)
Type	Self-employed	Wage-Earner	Self-employed	Self-employed	Wage-earner
Income Type	AGI	AGI	AGI	AGI	AGI
Years	1988-2006	1988-2006	1988-1993	1988-1993	1988-1993
δ_b	1,000	1,000	2,000	1,000	3,000
δ_c	1,000	1,000	1,000	1,000	1,000

The bunching estimates are calculated by first calculating the amount of imperfect bunching around the tax kink using a local linear regression as described in the text and the corresponding figures. Then, equation (1) is used to construct elasticities based on this estimate. The standard errors are calculated via non-parametric bootstrap and clustered by state (1000 replications). These estimates exclude individuals that did not file for the EITC, had no dependents, or had more than two dependents. The wage-earner estimates include only wage-earners and the self-employed include only self-employed individuals. $2\delta_b$ is the width of the imperfect bunching region and δ_c is the width of the counterfactual region.