Fiscal Policy and Inequality

Bunching Methods to Estimate Elasticities

Elliott Ash

ETH Zurich

Monday 29th October, 2018

Intro: Bunching Methods

- Tax changes often create discontinuities in the budget set:
  - **"Kinks"** and **"Notches"**
- How can we obtain elasticity estimates from individual responses to kinks and notches?
  - Main reference: overview paper by Kleven (2016)*
- Potential applications to other contexts in which there are cutoffs in the budget set:
  - Regulations (eg, labor laws)
  - Nonlinear pricing (eg, electricity consumption)

Bunching Method: Kinks

- In most settings, tax schedules are piece-wise linear due to progressivity
- This creates “kinks” in the budget set:
  - Kink = income threshold at which the **marginal** tax rate changes discontinuously
    - Discrete change in the slope of the budget set
    - The *average* tax rate does not jump – it changes continuously
- Kinks create endogeneity problem in regression analysis, because $\text{cov}(\tau, u) \neq 0$
- We can take advantage for **non-parametric identification**
  - Exploit response to kink to estimate behavioral elasticity

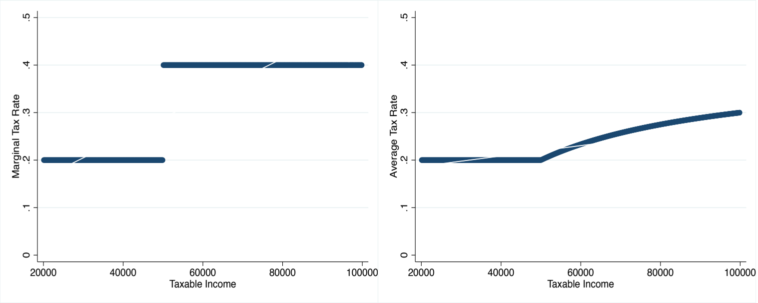
Bunching Method: Notches

- In some cases, taxes or regulations generate discontinuities in the budget set itself
- This creates a "notch":
  - Notch = income threshold at which the **average** tax rate changes discontinuously
    - Discrete change in the level of the budget set
    - The *marginal* tax rate may change as well
- Notches generate much stronger incentives than kinks
  - Often, create a "dominated" range where individuals should not locate under any utility function
  - Allows us to estimate the relevance of optimization frictions

Example: Kink in the tax schedule

- Consider a simple tax schedule:

$$\tau \equiv T'(z) = \begin{cases} 0.2 & \text{if } z \leq 50,000 \\ 0.4 & \text{if } z > 50,000 \end{cases}$$



Bunching: Relationship with other Methods

- Bunching shares common features with regression discontinuity (RDD) and regression kink (RKD) designs
  - In all cases, there is a threshold that determines different incentives
- Key difference:
  - Main assumption in RDD/RKD: agents cannot manipulate the running variable
    - ie, cannot choose what side of the threshold they are on
  - Bunching assumes that there is a response: we use bunching to estimate the underlying elasticity of *that* response

Bunching at Kinks: Setup

- **Key idea**: standard model predicts that individuals bunch at kinks in the budget set
- Start with linear tax system:
  - Marginal tax rate is τ for everyone, $T(z) = \tau z$
  - Utility: $u\left(c, \frac{z}{n}\right)$, where $c =$ consumption, $z =$ earnings, $n =$ ability
    - Assume no savings, so that $c \equiv z - T(z)$
    - $u\left(c, \frac{z}{n}\right) = u\left((1-\tau)z, \frac{z}{n}\right)$
  - Optimal choice determined by:

$$-\frac{u_c\left(c, \frac{z}{n}\right)}{u_z\left(c, \frac{z}{n}\right)} = (1-\tau)$$

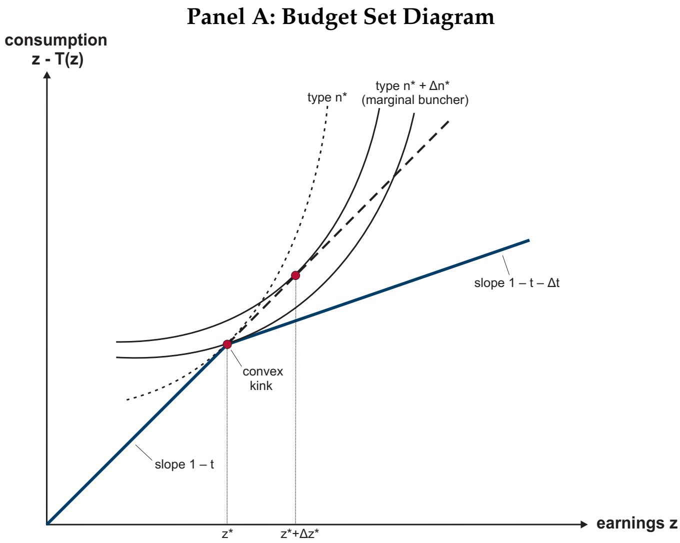
Small Kink Analysis

- Tax reform: new marginal tax rate $\tau + \Delta\tau$ (where $\Delta\tau > 0$ is small) for earnings above $z > z^*$, such that

$$T(z) = \tau \cdot z + \Delta\tau \cdot (z - z^*) \cdot \mathbf{I}[z > z^*]$$

  - where $\mathbf{I}[\cdot]$ is the indicator function
- This creates a kink in the budget set at income level z^*
- Net-of-tax rate goes *down*: $(1 - \tau)$ to $(1 - \tau - \Delta\tau)$

Bunching at Kinks: Budget Set



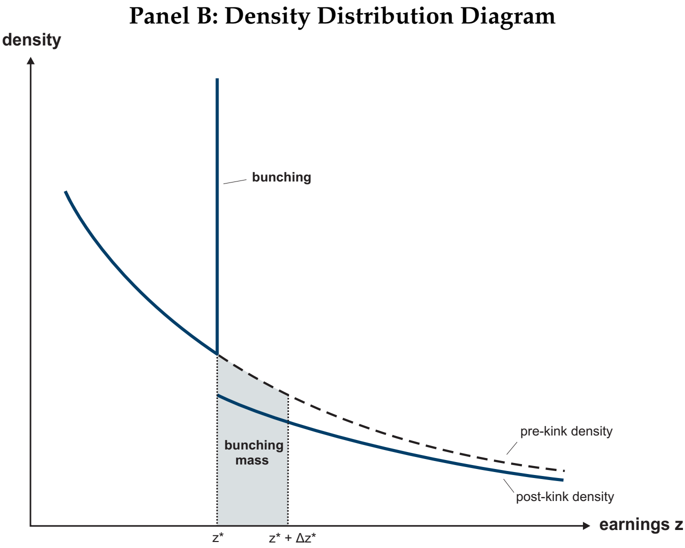
Source: Kleven (2015)

Bunching at Kinks: Budget Set

- With kink, indifference curves are tangent to different segments of the budget set
- Individual with ability n^* is unaffected
- Individual with ability $n \in (n^*, n^* + \Delta n^*$ reduces taxable earnings to z^* (ie, bunches at z^*):

Type (n)	Pre-reform	Post-reform
n^*	$z(n^*) = z^*$	$z(n^*) = z^*$
$n^* + \Delta n^*$	$z(n^* + \Delta n^*) = z^* + \Delta z^*$	$z(n^* + \Delta n^*) = z^*$

Bunching at Kinks: Density Distribution



Source: Kleven (2016)

Bunching at Kinks: Three groups

1. Unaffected by the change:
   - All individuals with $n \leq n^*$ are not affected, so their choice of z doesn't change
2. Bunchers:
   - Individuals with $n \in (n^*, n^* + \Delta n^*)$ will bunch at z^* after the reform
3. Adjusters:
   - Individuals with $n > n^* + \Delta n^*$ will reduce earnings after the reform, but they will locate above the kink (i.e., $z > z^*$)

Earnings Distribution Before/After Kink

- Before tax reform, z is distributed following $h_0(z)$, a smooth density function
  - All variation in $h_0(z)$ is due to differences in ability
- All individuals <u>initially</u> located between z^* and $z^* + \Delta z^*$ now bunch at the kink
  - Individual with type $n^* + \Delta n^*$ is the "marginal buncher"
- Number of **excess bunchers** is given by:

$$\begin{aligned} B &= \int_{z^*}^{z^*+\Delta z^*} h_0(z)\, dz \\ &\simeq h_0(z^*)\, \Delta z^* \end{aligned}$$

  - $h_0(z^*)$ = height of the pre-reform density at the kink
  - Implicit assumption: $h_0(z)$ is smooth around z^*

Estimating ETI using Kinks

- Definition of elasticity of taxable income (ETI):

$$\varepsilon(z) = \frac{\left[\frac{\Delta z}{z}\right]}{\left[\frac{\Delta(1-\tau)}{1-\tau}\right]}$$

- Evaluate at $z = z^*$ (kink):

$$\varepsilon(z^*) = \frac{\left[\frac{\Delta z^*}{z^*}\right]}{\left[\frac{\Delta(1-\tau)}{1-\tau}\right]}$$

- Use definition of $B \simeq h_0\left(z^*\right)\Delta z^*$, sub for Δz^*:

$$\varepsilon(z^*) = \frac{\left[\frac{B}{h_0(z^*)}\frac{1}{z^*}\right]}{\left[\frac{\Delta(1-\tau)}{(1-\tau)}\right]}$$

Formula for estimating ETI using kink

$$\varepsilon(z^*) = \frac{\left[\frac{b}{z^*}\right]}{\left[\frac{\Delta(1-\tau)}{(1-\tau)}\right]}$$

- $b \equiv \frac{B}{h_0(z^*)}$ = ratio of "excess bunchers" compared to counterfactual (pre-reform) density at the kink
- z^* = \$ income value where kink is located (see note in next slide)
- $\left[\frac{\Delta(1-\tau)}{(1-\tau)}\right]$ = % change in net-of-tax rate

$\Rightarrow$ Elasticity of Taxable Income (ε) is proportional to b, which can be easily estimated non-parametrically

ETI with Kinks: Discretize data

$z^* =$ \$ income value where kink is located

- Step 1:
  1. Discretize data into income bins (e.g., bins of \$1,000)
  2. Count taxpayers c_j in each bin j
- In the ETI formula, the value for z^* is the bin including, or just above, the kink

ETI with kinks: Form counterfactual

- To estimate B, we need to approximate the **counterfactual density** that would be observed in the absence of the kink
- General method: fitted polynomial regression

$$c_j = \underbrace{\sum_{i=0}^{p} \beta_i \cdot (z_j)^i}_{\text{polynomial}} + \underbrace{\sum_{i=z_-}^{z_+} \gamma_i \cdot \mathbf{I}\left[z_j = i\right]}_{\text{excluded range}} + \underbrace{v_j}_{\text{error term}}$$

  - c_j = number of individuals in bin j
  - p = order of the polynomial
  - $[z_-, z_+]$ = excluded range (a few bins above and below the kink)
    - this is equivalent to removing these bins from the data when running the regression.

ETI with kinks: Measure Bunching

- $\hat{c}_j$, predicted count of individuals by bin with fitted polynomial:

$$\hat{c}_j = \sum_{i=0}^{p} \hat{\beta}_i \cdot (z_j)^i$$

- Bunching is measured as the actual number of individuals at the kink, divided by the predicted counterfactual

$$\hat{b} = \frac{\sum_{j \in [z_-, z_+]} c_j}{\sum_{j \in [z_-, z_+]} \hat{c}_j}$$

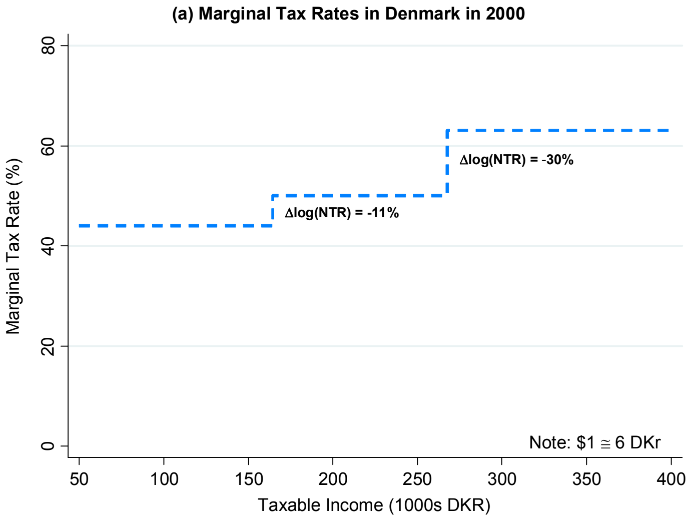
- Then the elasticity estimate is

$$\hat{\varepsilon}(z^*) = \frac{\left[\frac{\hat{b}}{z^*}\right]}{\left[\frac{\Delta(1-\tau)}{(1-\tau)}\right]}$$

Chetty et al. (QJE, 2011): Bunching in Denmark

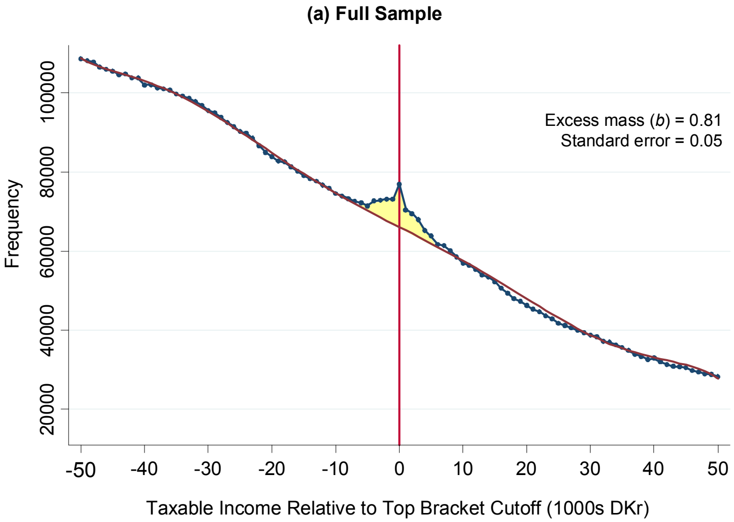
- Study taxpayer behavior under Danish income tax
- Marginal income tax rate goes up from **49% to 63%** for earnings above DKr 267,600 (£30,000)
  - Large kink: $(1 - t)$ goes down from 0.51 to 0.37
  - 30% fall in net-of-tax rate
- **Administrative Data**: universe of Danish taxpayers, 1994-2001
  - About 4 million taxpayers per year

Danish Income Tax



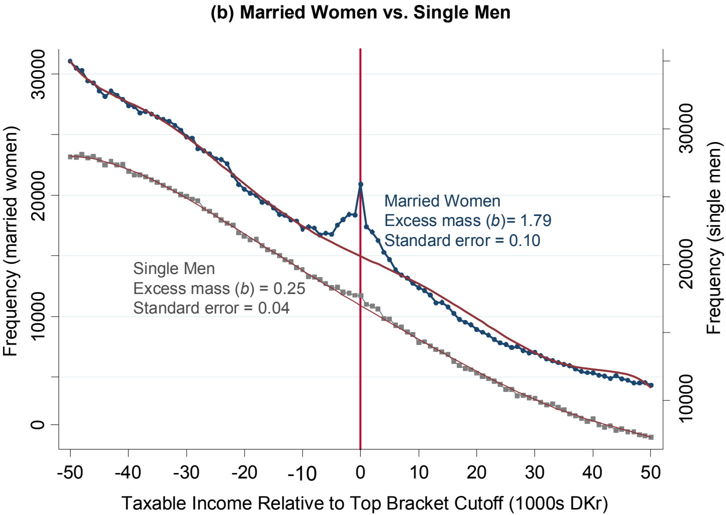
Chetty et al. (QJE, 2011): Results

Full Sample of Individual Taxpayers



Chetty et al. (QJE, 2011): Results

Married women vs. Single Men



Chetty et al. (QJE, 2011): Results

- Bunching response very strong for married women
  - Often second earners in the household
- Weak response for single men
  - Little flexibility in labor supply

Chetty et al. (QJE, 2011): Results

TABLE II
Observed Elasticity Estimates using Small Tax Reforms

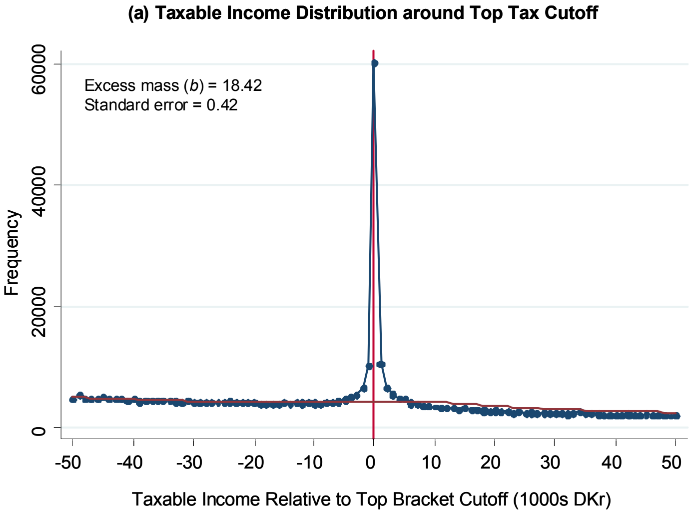
	Dependent Variable: Log Change in Wage Earnings					
Subgroup:	All Wage Earners		Married Females	High-Experience Married Female Professionals	Wage Earners 100-300K	Wage Earners > 200K
	(1)	(2)	(3)	(4)	(5)	(6)
log change in net-of-tax rate (Δ log (1-t))	-0.001 (0.003)	-0.004 (0.003)	0.006 (0.005)	0.000 (0.011)	-0.006 (0.003)	-0.001 (0.003)

Source: Chetty, Friedman, Olsen & Pistaferri (QJE, 2011)

Chetty et al. (QJE, 2011): Results

- Very small elasticities for wage earners: $\varepsilon \approx 0$
- Overall, surprisingly small elasticity estimates
  - Compare with Kleven and Schultz (2014) estimates, obtained using DD methods with exactly the same data and similar tax reforms

Bunching among Self-Employed

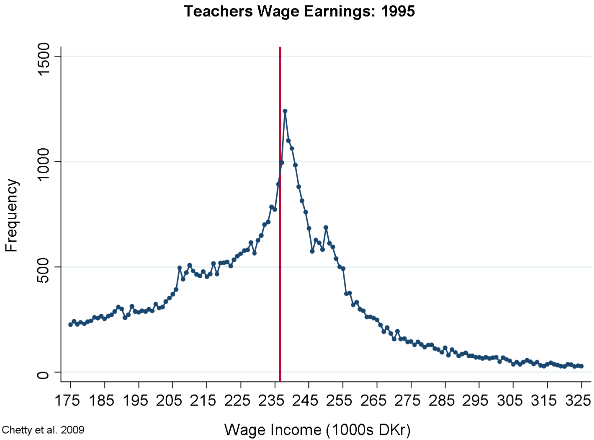


Chetty et al. (QJE, 2011): Self-Employed

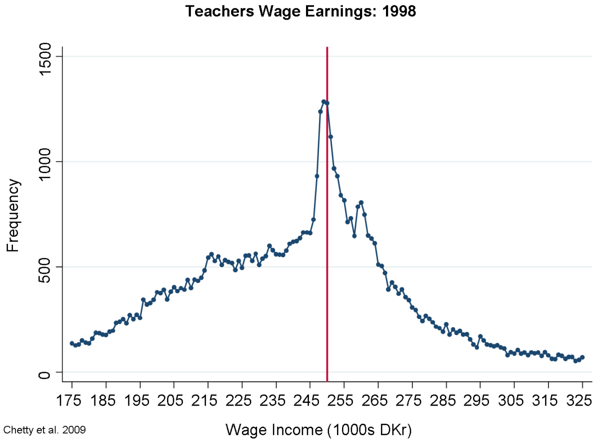
- Higher elasticity for self-employed: $\varepsilon = 0.24$
  - Self-employed bunch very strongly. Why?

1. More flexible labor supply (**real** response)
   - Can adjust hours much more easily
2. Higher ability to manipulate reported income (**avoidance/evasion**)
   - Harder to track payments because there is no third-party reporting

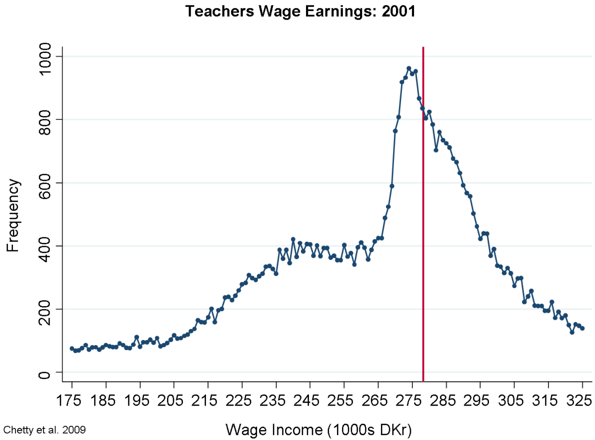
Bunching among Teachers, 1995



Bunching among Teachers, 1998



Bunching among Teachers, 2001



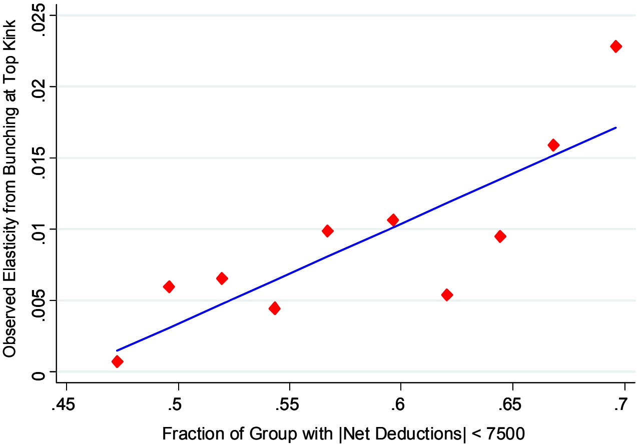
Chetty et al. (QJE, 2011): Teachers

- Teachers' salaries bunch very strongly. Why?
  - Little flexibility in labor supply
  - Unlikely to know details of tax system
- But: Strong teachers' unions negotiate salaries every year!
  - They follow changes in tax system
  - Everyone in the union benefits

Chetty et al. (QJE, 2011): Results

Elasticity correlated with Number of Available Deductions

FIGURE X
Observed Elasticities vs. Scope of Tax Changes



Chetty et al. (QJE, 2011): Results

- Overall, surprisingly small elasticity estimates. Why?

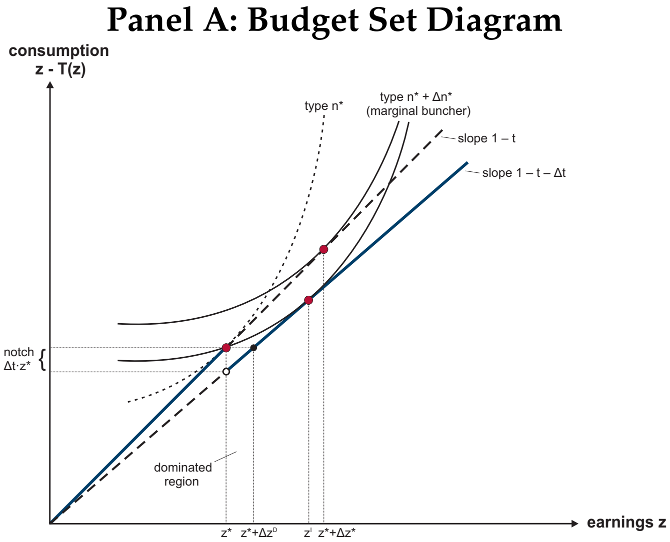
1. Adjustment costs
    1.1 Cost of re-optimizing may be higher than benefit
2. Constraints on hours of work per week
3. Inattention?

Notched Income Tax Schedule in Pakistan

Average tax rates

Source: Kleven and Waseem (QJE, 2013)

Notched Budget Set



Source: Kleven (2015)

Notched Budget Set

- **Average** tax rate jumps, instead of *marginal* tax rate
  - Much stronger incentives: £1 in additional earnings can lead to ***lower*** after-tax income
- Tax schedule given by

$$T(z) = \tau \cdot z + \Delta\tau \cdot z \cdot \mathbf{I}[z > z^*]$$

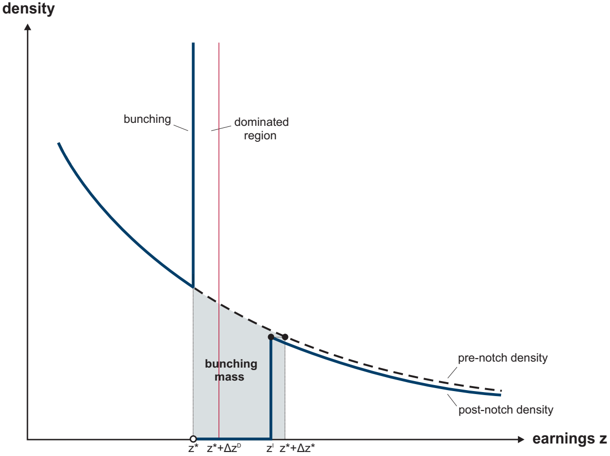
  - where τ is the marginal tax rate
  - **Note**: τ is the average tax rate in the first bracket, and $\tau + \Delta\tau$ is the average tax rate in the second bracket
- With notches, bunching should be observed to the left of the notch
  - With kinks, bunching *around* the threshold. Why?

Bunching at Notches: Dominated Range

- Notches create a “dominated range”:
  - Irrational to locate in $\left(z^*, z^* + \Delta z^D\right)$ under *any* preferences
  - We should expect a “hole” in the distribution above the notch...
  - ...unless there are optimization frictions

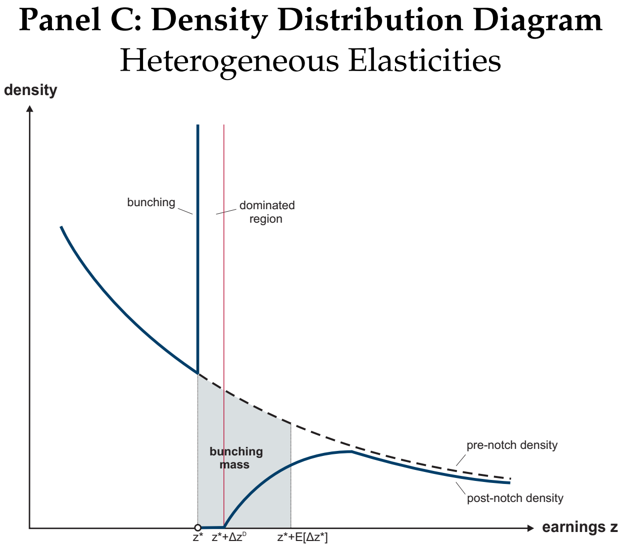
Bunching at Notches: Homogeneous Elasticities

Panel B: Density Distribution Diagram
Homogeneous Elasticities



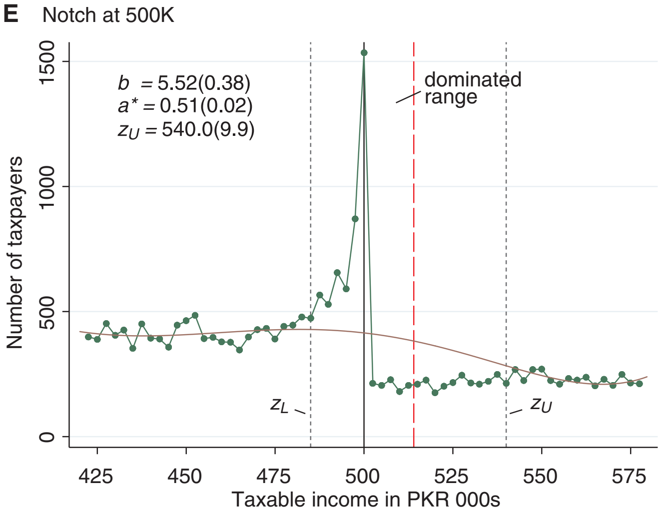
Source: Kleven (2015)

Bunching at Notches: Heterogeneous Elasticities



Source: Kleven (2015)

Kleven & Waseem (QJE, 2013): Bunching at a Notch



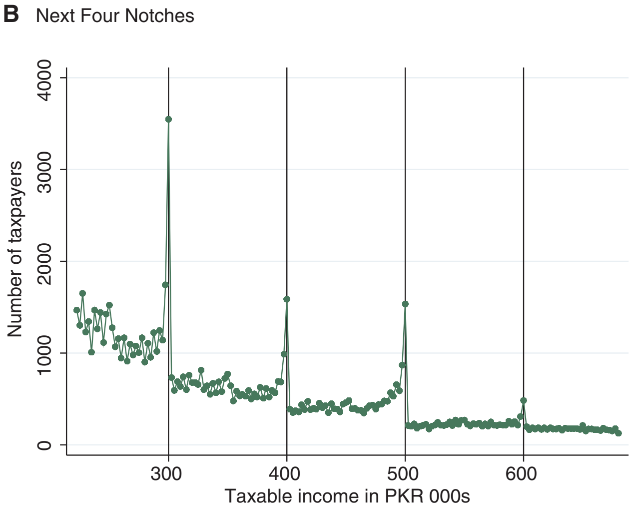
Quantifying Frictions with notches

- With notches, we can leverage the dominated range to quantify the importance of **optimization frictions**
- In theory, we should observe a hole in $\left[z^*, z^* + \Delta z^D\right]$
  - But we don't in practice. How to allow for that?
- Let a proportion a^* of individuals be in the dominated range, we interpret that a^*% of individuals are affected by frictions
- Then, we reweight our bunching estimates by a factor $(1 - a^*)$
  - Similar to treatment-effect-on-the-treated (ToT) estimates

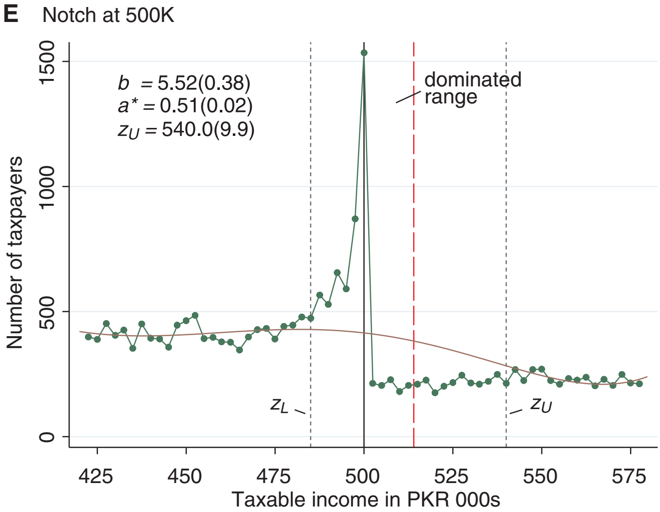
Kleven & Waseem (QJE, 2013): Income Tax in Pakistan

Average tax rates

Kleven & Waseem (QJE, 2013): Income Distribution



Kleven & Waseem (QJE, 2013): Bunching at a Notch



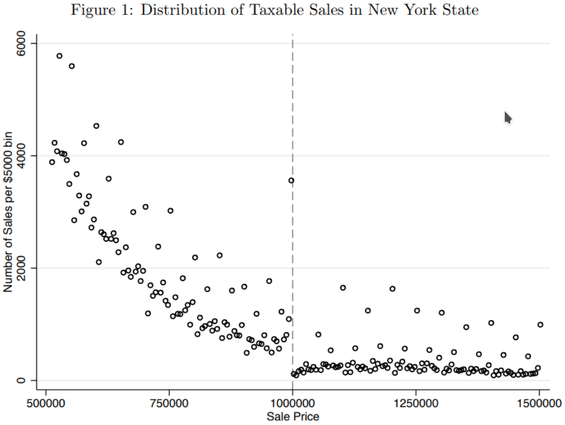
Kleven & Waseem (QJE, 2013): Results

- Taxpayers bunch strongly at each notch
- But there are some taxpayers in the dominated range. Why?
  - Optimization frictions: adjustment costs and inattention
  - Career concerns: current earnings may affect future earnings

Kleven & Waseem (QJE, 2013): Results

- Generally low elasticities ($\epsilon \approx .02$), despite large bunching response
- Elasticities larger for self-employed ($\epsilon \approx .15$), as in other contexts
- Optimization frictions are very important:
  - Despite strong incentives created by notches, many people do not modify their economic decisions

Distribution of Housing Sale Prices, New York State



- Kopcuk and Munroe (2014)