IDENTIFICATION IN BUNCHING DESIGNS

Henrik Kleven

Princeton University and NBER

March 2018

Bunching Equation

$$B = \int_{y^*}^{y^*+\Delta y^*} f_c\left(y\right) dy$$

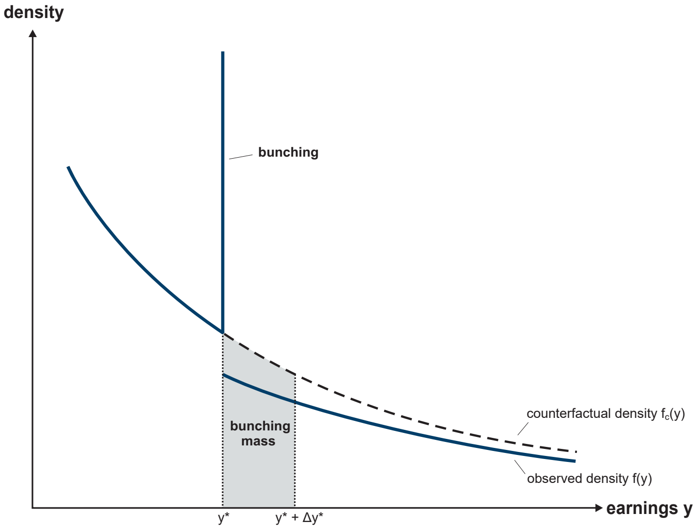
- B is bunching , $f_c\left(y\right)$ is a counterfactual density for outcome y, y^* is the threshold, Δy^* is the response
- The idea is simple:
  Estimate counterfactual $f_c(y) \rightarrow$ bunching $B \rightarrow$ response Δy^*

Identification Challenge I: Counterfactual Density

$$B = \int_{y^*}^{y^*+\Delta y^*} f_c\left(y\right) dy$$

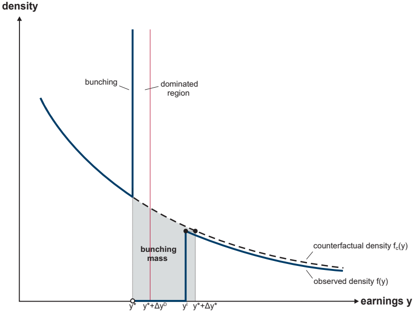
- The approach relies on estimating $f_c(y)$ on a discrete and *a priori* unknown bunching segment $(y^*, y^* + \Delta y^*)$

Bunching in Theory: Kinks

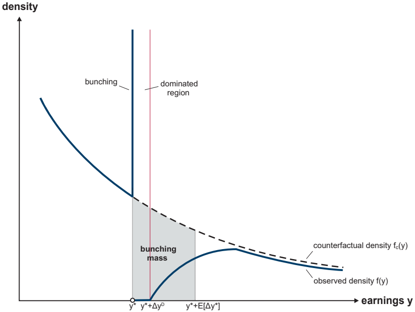


Bunching in Theory: Notches

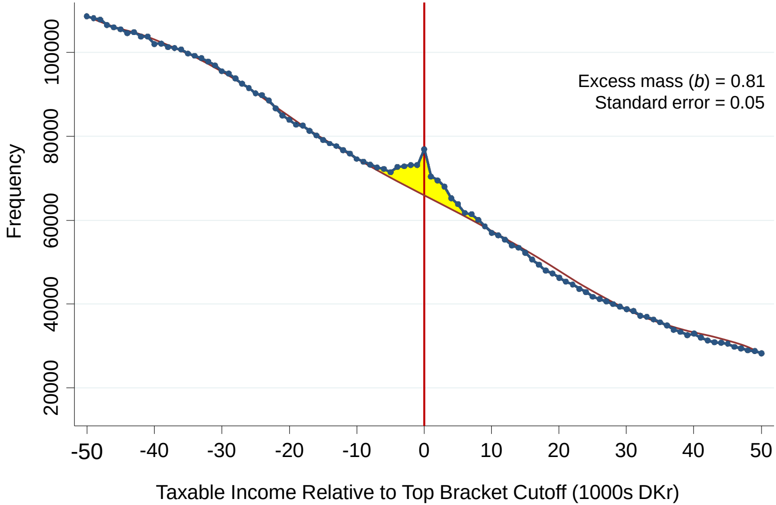
Homogeneous Elasticities



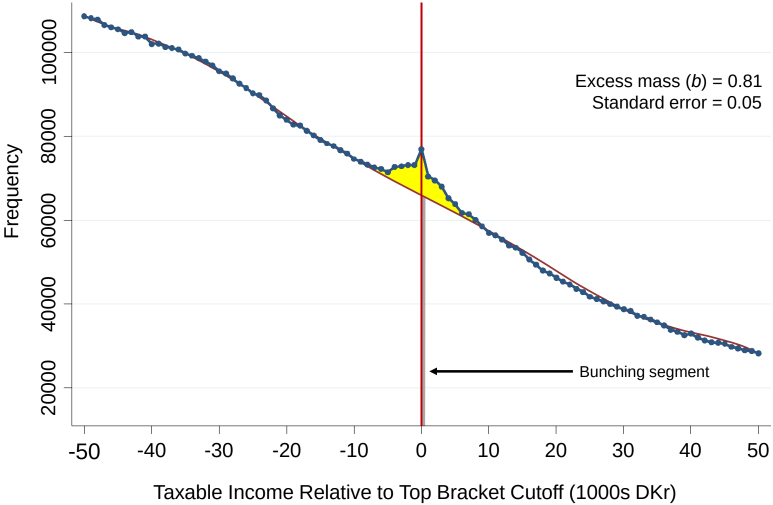
Heterogeneous Elasticities



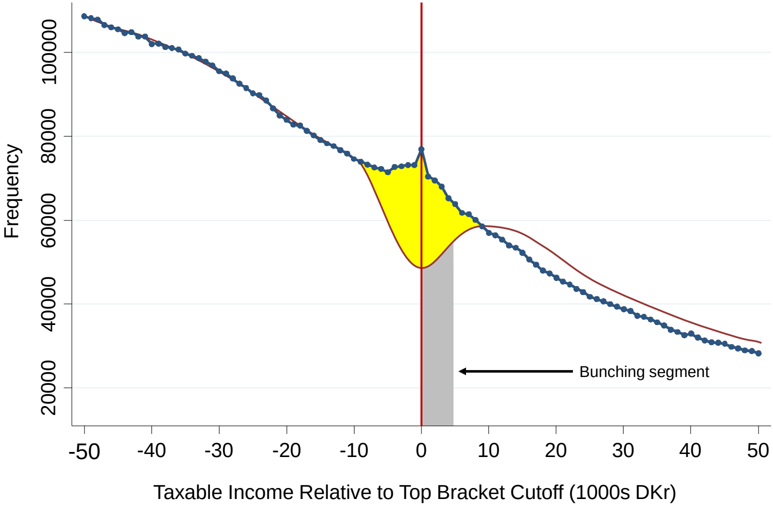
Bunching in Reality (Chetty et al. 2011): Bunching is Diffuse



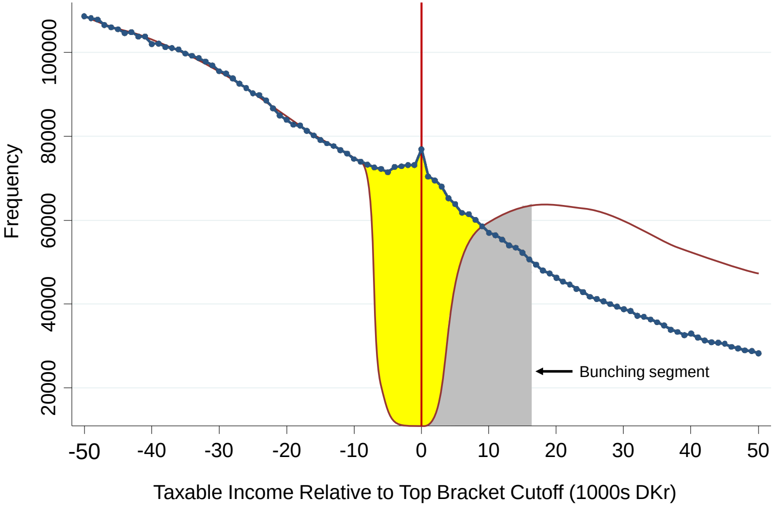
Bunching in Reality (Chetty et al. 2011): Bunching Segment is Tiny



Bunching in Reality (Chetty et al. 2011): But We Could Consider Another Counterfactual



Bunching in Reality (Chetty et al. 2011): Or Something Even Crazier



Identification Challenge I: Counterfactual Density

$$B = \int_{y^*}^{y^*+\Delta y^*} f_c(y)\, dy$$

- Assuming only smoothness (f_c is $\mathbf{C^1}$), it is mathematically possible to get any Δy^*
  - Without any additional assumptions or moments, we cannot be sure to what degree bunching reflects a behavioral response Δy^* or just an unusual distribution of preference heterogeneity
  - Formalized by Blomquist & Newey (2017)
- This point is obvious and already understood, but clarifying language may be useful

Identification Challenge I: Counterfactual Density

$$B = \int_{y^*}^{y^* + \Delta y^*} f_c(y)\, dy$$

- Whether the estimation is compelling depends on the data
  - E.g., the counterfactual in Chetty et al. (2011) is compelling
- What's more, we have focused on the worst-case scenario:
  - One cross-section of data around one kink $\rightarrow$ counterfactual must come from extrapolating the observed distribution
  - When this is not compelling, additional variation can be used: temporal variation in kink, cross-sectional variation in kink, panel data (see Kleven 2016)

Estimating the Counterfactual Buys Us Δy^*

$$B = \int_{y^*}^{y^* + \Delta y^*} f_c(y)\, dy$$

- But Δy^* is not an interesting, externally valid entity
  - E.g., the estimate of Δy^* in Chetty et al. (2011) tells us that the (average) marginal buncher moves down 810 DKK. So What?

- We have not yet identified a parameter of interest
  $\rightarrow$ We need a model

Identification Challenge II: Relating Δy^* to Elasticity e

$$B = \int_{y*}^{y^* + g(e, \boldsymbol{x})} f_c(y)\, dy$$

- Saez (2010): $\Delta y^* = g(e, \boldsymbol{x})$ where e is the Hicksian labor supply elasticity and $\boldsymbol{x}$ is a vector of observables
- This requires a very specific model:
  - Static, frictionless, deterministic environment with perfect compliance and no behavioral biases

Identification Challenge II: It Gets Much Harder

$$B = \int_{y^*}^{y^* + g(e, \mathbf{\Phi}, \boldsymbol{x})} f_c(y)\, dy$$

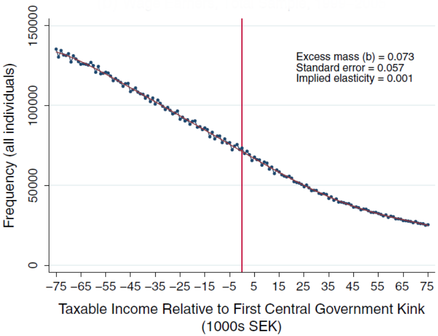
- We have $\Delta y^* = g(e, \mathbf{\Phi}, \boldsymbol{x})$ where $\mathbf{\Phi}$ includes factors that affect Δy^* for a given Hicksian elasticity e
  - Without assumptions or evidence on $\mathbf{\Phi}$, a bunching response Δy^* is consistent with *any elasticity between zero and infinity*
  - This is the real identification challenge (see Kleven 2016)

Identification Challenge II: What Goes Into Φ?

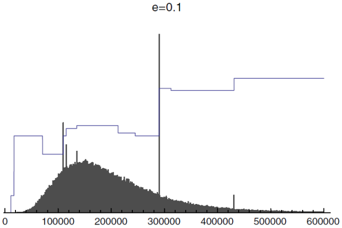
- Evasion and avoidance
- Dynamic aspects (intertemporal substitution, career effects)
- Uncertainty (random income components)
- Lumpiness (indivisibility of hours)
- Adjustment costs (search costs)
- Inattention and misperception
- Reference dependence (round numbers, policy-induced focal points)

Sometimes We See Zero Bunching At Huge Kinks
Bastani & Selin (2014)

Observed Bunch

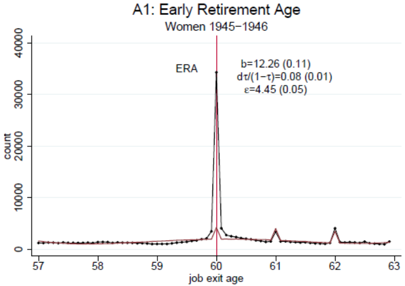


Simulated Bunch Under $e = 0.1$

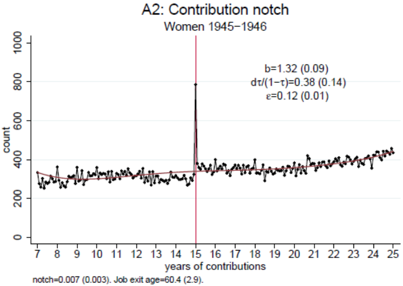


Sometimes We See Huge Bunching At Tiny Kinks
Seibold (2018)

Huge Bunch at Tiny Kink



Smaller Bunch at Notch



Kinks vs Notches (Saez vs Kleven)

- Two basic differences:
  - Notches create much larger incentives
  - Kinks are common in personal income taxation (ETI estimation), but notches are ubiquitous more broadly
- *In practice*, these strategies pose a trade-off between the two identification challenges I have discussed
  - Estimating the counterfactual can be harder with notches
  - Relating Δy^* to a structural elasticity is easier with notches:
    - Incentives are larger $\rightarrow$ optimization frictions less of an issue
    - Strictly dominated regions can be used to measure friction

We're Learning From the Bunching Literature, But Not the Labor Supply Elasticity

- Evidence that there *is* a response (non-trivial in some contexts)
- Evidence on other (more frictionless) outcomes
  - Evasion/avoidance is one, but many other (non-tax) outcomes
- Learn about optimization frictions
- Learn about reference dependence
- Learn from the *variation* in bunching by another dimension x, leveraging observational or experimental variation in x
- Use bunching as an "instrument" for something else
- Developing models that can predict bunching behavior is challenging and has broader implications for understanding behavior