Bunching

Shanon Hsuan-Ming Hsu

University of Chicago

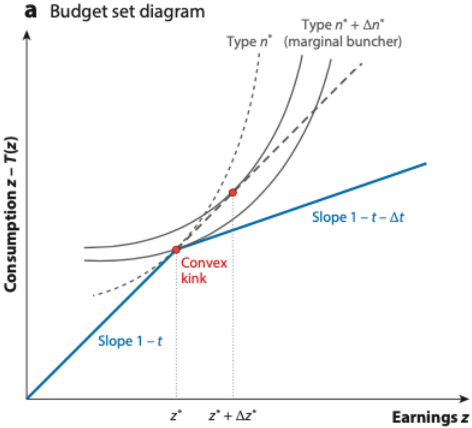
May 21, 2020

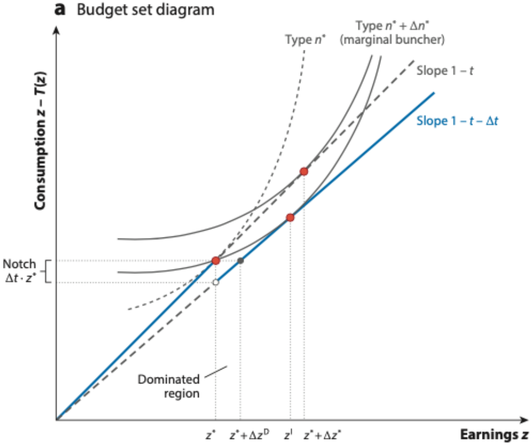
Overview

- The idea is to exploit bunching around points that feature discontinuities in incentives to identify and estimate structural parameters
- Two conceptually different bunching designs:
  - Kink points: discrete changes in the **slope** of choice sets (Saez, 2010; Chetty et al., 2011)
  - Notch points: discrete changes in the **level** of choice sets (Henrik J Kleven and Waseem, 2013)
- Early literature focuses on estimating labor supply elasticity using kinked budget sets
  - The insight is that the more the bunchers, the larger the elasticity
- Recent applications in other settings: electricity demand, attribute-based regulations, prescription drug insurance, labor regulations, education policy, etc
  - See Henrik Jacobsen Kleven (2016) and Bertanha, McCallum, and Seegert (2019) for more

Illustration of kink and notch designs

Figure: Kink and notch





Classic example income tax and labor supply

Saez (2010) and Henrik J Kleven and Waseem (2013)

- Individual with heterogenous ability n chooses labor income z and consumption $c = z - T(z)$:

$$\max_{c,z} \; U(c, z) = \underbrace{z - T(z)}_{c} - \frac{n}{1 + 1/\varepsilon} \left(\frac{z}{n} \right)^{1+\frac{1}{\varepsilon}},$$

where there is a linear tax system with a kink at z^*:

$$T(z) = \mathbf{1}\{z \leq z^*\}\, t_0 z + \mathbf{1}\{z > z^*\}\, t_1 z,$$

with $t_0 < t_1$ and $\Delta t \equiv t_1 - t_0 > 0$.

Solution to UMP

- The solution to the utility maximization problem for agent of ability n is

$$z = \begin{cases} n(1-t_0)^{\varepsilon} & \text{if } n < \underline{n} \\ z^* & \text{if } \underline{n} \leq n < \overline{n} \\ n(1-t_1)^{\varepsilon} & \text{if } \overline{n} < n, \end{cases}; \quad \begin{cases} \underline{n} = z^*(1-t_0)^{-\varepsilon} \\ \overline{n} = z^*(1-t_1)^{-\varepsilon} \end{cases}$$

- Bunching is defined as:

$$B \equiv \mathbb{P}(z = z^*) = \mathbb{P}(\underline{n} \leq n \leq \overline{n}) = \int_{\underline{n}}^{\overline{n}} f_n(u)du = F_n(\overline{n}) - F_n(\underline{n})$$

$$= \mathbb{P}(z^* \leq z \leq z^* + \Delta z^*) = \int_{z^*}^{z^*+\Delta z^*} b(z)dz$$

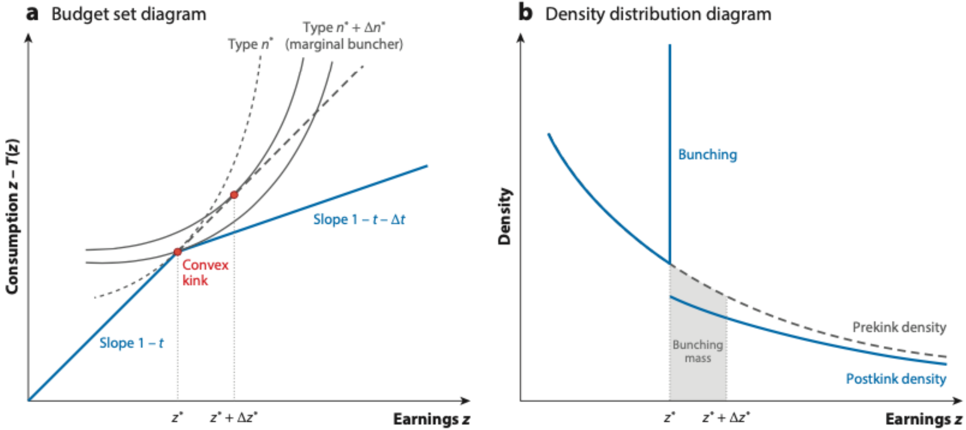
  - the marginal buncher of ability $\overline{n}$ had there been no kink would choose:

$$\overline{n}(1-t_0)^{\varepsilon} = z^*(1-t_1)^{-\varepsilon}(1-t_0)^{\varepsilon} = z^*\left(\frac{1-t_0}{1-t_1}\right)^{\varepsilon} \equiv z^* + \Delta z^*$$

  - $b(z)$ is the counterfactual distribution of earning had there been no kink

Bunching diagram

Figure: Kink analysis



Identification of labor supply/earning elasticity

- One can solve the earnings elasticity ε as a function of Δz^*

$$\varepsilon = \frac{\log\left(1 + \Delta z^*/z^*\right)}{\log\left(1 - t_0/(1 - t_1)\right)}$$

- Saez (2010) uses a trapezoidal approximation to $b(z)$ around the kink to obtain

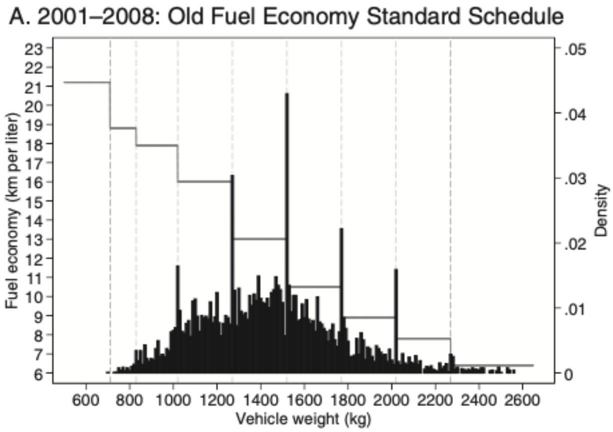
$$B = \int_{z^*}^{z^*+\Delta z^*} b_z(z)dz \approx \frac{[b(z^* + \Delta z^*) + b(z^*)] \cdot \Delta z^*}{2}$$

  - $b(z^*) = \tilde{b}(z^*)_-$, $b(z^* + \Delta z^*) = \tilde{b}\left(z^*\right)_+ \left(\frac{1-t_1}{1-t_0}\right)^{\varepsilon}$, and bunching mass B observed
  - Thus, Δz^* is identified and so is ε
- Bertanha, McCallum, and Seegert (2019) points out that this type of approximation may not be innocuous: underlying assumption on the shape of unobserved $b(z)$
  - non-parametric identification of ε impossible with only kink (but possible with notch)
  - partial identification with shape restrictions

Estimation of bunching with polynomial strategy

Notch points in Ito and Sallee (2018)

Figure: Fuel Economy Standard and Histogram of Vehicles Weights



Estimate the counterfactual distribution and bunching

Notch points in Ito and Sallee (2018)

- Grouping cars into small-weight bins (e.g., 10 kg) and denote the number of cars in bin j by c_j with average weight w_j

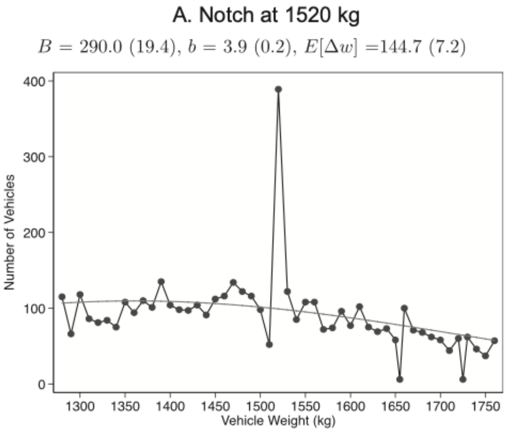
$$c_j = \sum_{s=0}^{S} \beta_s^0 \times (w_j)^s + \sum_{k=1}^{K} \gamma_k^0 \times d_k + \varepsilon_j$$

  - β_s^0 is initial estimate for polynomial fit; γ_k^0 captures bin fixed effect for notches $k = 1, \ldots, K$
  - Counterfactual distribution had there been no bunching is $\hat{c}_j^0 = \sum_{s=0}^{S} \beta_s^0 \cdot (w_j)^s$
  - Bunching estimate at notch k is given by $\hat{B}_k^0 \equiv c_k - \hat{c}_k$
- Problem: the area under the counterfactual distribution not equal the area under the empirical distribution
  - That is, the counterfactual p.d.f. is not a probability distribution as it does not integrate to 1
  - Need to "*shift*" the counterfactual distribution upward until it the integration constraint holds

Fitting polynomial to counterfactual

Notch points in Ito and Sallee (2018)

Figure: Bunching estimation with polynomial fit



Shifting the counterfactual distribution

Notch points in Ito and Sallee (2018)

- Shifting the counterfactual distribution relies on knowledge of where the excess bunching comes from
  - In this context, bunching mass comes from the left of the weight distribution
  - In Saez (2010), bunching mass comes from the right of the earnings distribution
- Parametric assumption about the distribution of bunching: shifting the counterfactual distribution uniformly from the range of $j \in (k-1, k)$:

$$c_j + \sum_{k=1}^{K} \alpha_{kj} \times \hat{B}_k = \sum_{s=0}^{S} \beta_s \times (w_j)^s + \sum_{k=1}^{K} \gamma_k \times d_k + \varepsilon_j \qquad (1)$$

where

$$\alpha_{kj} = \begin{cases} \frac{c_j}{\sum_{j \in (k-1,k)} c_j} & \text{for} j \in (k-1, k) \\ 0 & \text{otherwise.} \end{cases}$$

- Since $\hat{B}_k = c_k - \hat{c}_k = c_k - \sum_{s=0}^{S} \hat{\beta}_s \cdot (w_j)^s$ is a function of $\hat{\beta}_s$, estimate (1) by iteration until reaching a fixed point

Other identification assumptions/issues

Following Henrik Jacobsen Kleven (2016)

- Smoothness of the counterfactual distribution
  - If other policies change at the same threshold, then the distribution may not be smooth
- Shape of counterfactual (locally), as discussed
- Heterogeneous elasticity
  - in this case, bunching identifies the average response across ε_i's
- Measurement error and optimization friction
  - Only $\tilde{z} \equiv z + e$ is observed where e is noise
  - Bunching is larger (in elasticity terms) when the kink/notch is larger

Reference

Bertanha, Marinho, Andrew H McCallum, and Nathan Seegert (2019). "Better bunching, nicer notching". In: *Nicer Notching (December 23, 2019)*.

Chetty, Raj et al. (2011). "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records". In: *The quarterly journal of economics* 126.2, pp. 749–804.

Ito, Koichiro and James M Sallee (2018). "The economics of attribute-based regulation: Theory and evidence from fuel economy standards". In: *Review of Economics and Statistics* 100.2, pp. 319–336.

Kleven, Henrik J and Mazhar Waseem (2013). "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan". In: *The Quarterly Journal of Economics* 128.2, pp. 669–723.

Kleven, Henrik Jacobsen (2016). "Bunching". In: *Annual Review of Economics* 8, pp. 435–464.

Saez, Emmanuel (2010). "Do taxpayers bunch at kink points?" In: *American economic Journal: economic policy* 2.3, pp. 180–212.