Bunching

Skand Goel

July 27, 2018

Key idea

- "Discontinuities" lead to corner solutions in several choice problems.
- These show up as bunched outcomes in choice data.
- Use this mapping to say something about behavioral response and/or structural parameters.
- Opposite of regression discontinuity.

Structure of talk

- Simple labor supply examples
- Some extensions (in the context of labor supply)
- Estimation
- Briefly, one application that is not labor supply.

Labor supply

- Agent's utility depends on consumption (after tax income) and effort (before tax income)

$$u\left(z - T(z), \underbrace{\frac{z}{n}}_{z=nl}\right)$$

  where

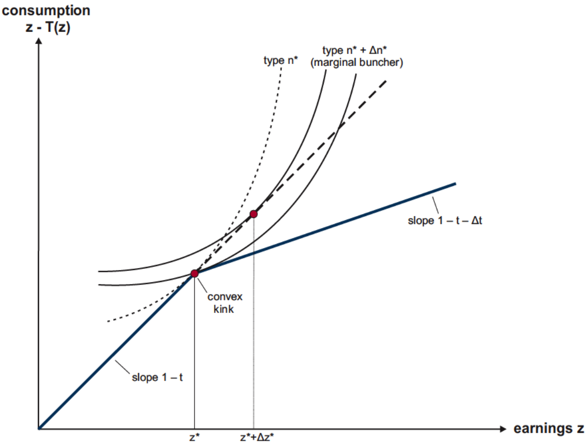
  - z: earnings (choice variable),
  - $T(z)$: Tax function,
  - n: ability (exogenous heterogeneity) with distribution $f(n)$.

- Baseline: With $T(z) = t.z$, we have a linear budget constraint.
- $z^{\text{opt}}(n)$ implies the (endogenous) density of z, $h_0(z)$.

Kink in budget constraint

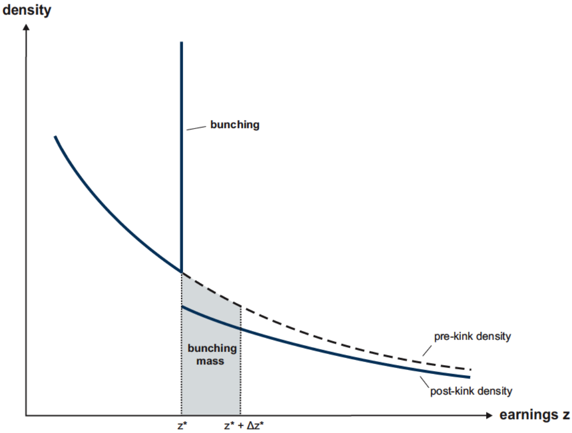
Consider a (small) change in the marginal tax rate at z^*

$$T(z) = tz + \triangle t.(z - z^*)\mathbf{I}(z > z^*)$$



Bunching in earnings density

Suppose data was generated from this model



Linking bunching to elasticity

- $\triangle z^*$ represents the interior earnings response between two tangency points.
- Define "elasticity" of earnings with respect to $(1-t)$

$$e = \frac{\triangle z^* / z^*}{\triangle t / (1-t)}$$

- Moreover, total bunching B related to e:

$$B = \int_{z^*}^{z^* + \triangle z^*} h_0(z) dz \approx h_0(z) \triangle z^*$$

assuming (for now) $h_0(z)$ is locally constant.

Identifying e

$$e = \frac{\triangle z^* / z^*}{\triangle t / (1 - t)}$$

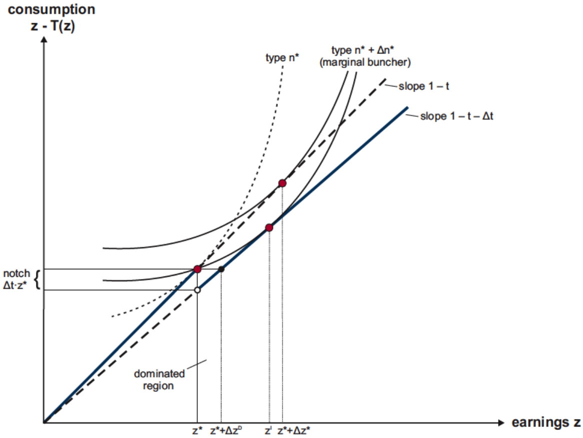
$$B \approx h_0(z) \triangle z^*$$

- $h_0(z)$ is counterfactual density, not observed.
- If we had an estimate of $h_0(z)$, we could back out $\triangle z^*$ and therefore e.

Notch in budget constraint

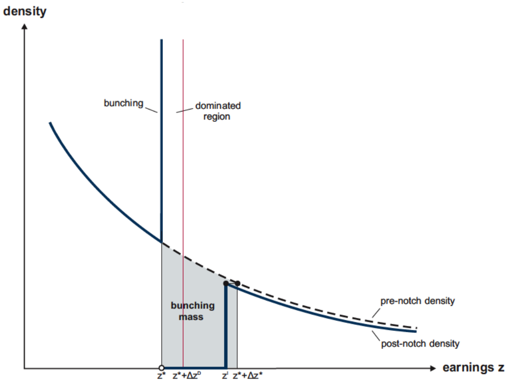
Consider a (small) change in the *average* tax rate at z^*

$$T(z) = t.z + \triangle t.z.\mathbf{I}(z > z^*)$$



Bunching and hole in earnings density

Suppose data was generated from this model



Linking bunching to elasticity

- Not straightforward because marginal tax rate t varies by z.
- $t^* \equiv T(z^* + \triangle z^*) - T(z^*) \approx t + \triangle t . z^* / \triangle z^*$
- Define elasticity of earnings with respect to $(1 - t)$

$$e = \frac{\triangle z^* / z^*}{\triangle t / (1 - t)} \approx \frac{(\triangle z^* / z^*)^2}{\triangle t / (1 - t)}$$

- Again, total bunching B related to e:

$$B = \int_{z^*}^{z^* + \triangle z^*} h_0(z) dz \approx h_0(z) \triangle z^*$$

assuming (for now) $h_0(z)$ is locally constant.

Overview of estimating $\hat{h}_0(z)$

- Data: Cross-section of earnings and tax rule.
- Fit a flexible polynomial to observed distribution, excluding a window around z^*. Then, extrapolate to z^*.
- Group data into bins j and estimate

$$c_j = \Sigma_{i=0}^{p} \beta_i \left(z_j\right)^i + \Sigma_{i=z^-}^{z^+} \gamma_i \mathbf{I}\left[z_j = i\right] + \nu_j$$

  where
  - c_j: number of people in bin j
  - $[z^-, z^+]$ is the excluded range and γ_i are dummies
- To extrapolate, omit γ_is.

Overview of estimating $\hat{h}_0(z)$

- This approach valid under the assumption that $h_0(z)$ is a smooth function.
- Order of polynomial amounts to a shape restriction
  - Required for identification
- Choices
  - Window around threshold $[z^-, z^+]$
  - Order of polynomial p

Next steps

- Extensions of the model
  - Structurally interpret e
  - Deal with inconsistencies between model and data
- Choosing estimation parameters for h_0

Compensated vs uncompensated elasticities

- $\triangle z^*$ is very small, there are no income effects on the marginal buncher. e is compensated elasticity.
- Else, it is a weighted average of compensated and uncompensated elasticities.
- $\triangle z^*$ depends on the size of the kink/notch.
- Alternative - start with a quasilinear utility function

$$u = z - T(z) - \frac{n}{1 + 1/e} \left(\frac{z}{n} \right)^{1+1/e}$$

Interpreting e: Kinks

- The marginal buncher is tangent to the new budget line at z^*:

$$z^* = (n + \triangle n^*)\,(1 - t - \triangle t)^e$$

- Also, tangent to counterfactual earnings at $z + \triangle z^*$:

$$z^* + \triangle z^* = (n + \triangle n^*)\,(1 - t)^e$$

- Together, we have the "correct" expression

$$e = -\frac{\log\left(1 + \triangle z^* / z^*\right)}{\log\left(1 - \triangle t / (1 - t)\right)}$$

- As before, we get $\triangle z^*$ from the equation for B.
- For small $\triangle t$ and, therefore, small $\triangle z^*$, this is the same as the reduced form expression.

Interpreting e: Notches

- Marginal buncher for notches defined by

$$u(z^*) = u(z^I)|_{\text{tangent to new budget}}$$

- This implicitly characterizes the relationship between $\triangle z^*/z^*$, $\triangle t/(1-t)$ and e.
- More general idea - model needed to map $\triangle z^*$ to structural e. e.g. dynamics

Heterogenous e

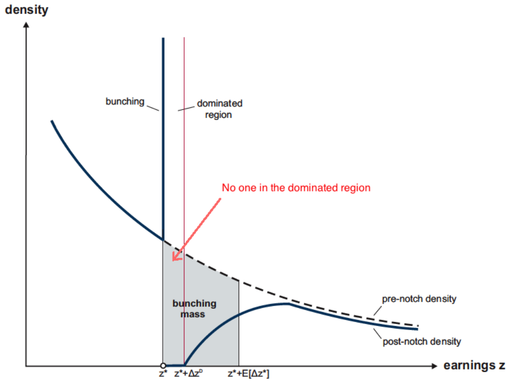
- The analysis till now assumed e was homogenous, but we can think of $\tilde{f}(n, e)$.
- Now, the relationship between B and $\triangle z_e^*$ becomes

$$B = \int_e \int_{z^*}^{z^* + \triangle z_e^*} \tilde{h}_0(z, e) \mathrm{d}z \mathrm{d}e \approx h_0(z^*) E\left[\triangle z_e^*\right]$$

- With an estimate of $h_0(z^*)$, can back out $E\left[\triangle z_e^*\right]$ and plug into some e formula.
- Aggregation bias: Elasticity of average response $\neq$ Average elasticity.
- However, can be bounded.

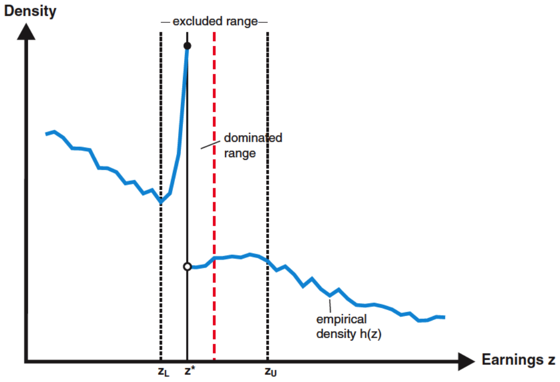
Heterogenous e

Implication for data: In case of notches, and smooth $\hat{f}(n, e)$, the hole no longer has a sharp edge:



Model vs. data

Notches lead to dominated region, but what if data didn't concur -



(At least) two tweaks to theory:

- Optimization frictions
- Reference points

Optimization frictions

- One way of rationalizing the data is that some agents do not optimize well.
- In the frictionless model, $B = B(e, \mathsf{x})$.
- With frictions, $B = B(e, \phi, \mathsf{x})$. ϕ parametrizes optimization frictions.
- If data were generated from a frictionless world, ϕ not identified.
- However, the type of data above generates an additional moment - the number of people in the dominated region.

Optimization frictions

- Let $a(z, e, \phi)$ denote share of (z, e) type individuals who are unresponsive. Then

$$B = \int_e \int_{z^*}^{z^* + \triangle z_e^*} (1 - a(z, e, \phi))\, \tilde{h}_0(z, e) dz de$$
$$\approx h_0(z^*) \left(1 - a^*(\phi)\right) E\left[\triangle z_e^*\right]$$

  where approximation now also assumes locally constant $a(z, e, \phi) = a^*(\phi)$.
- As $a^*(\phi)$ is observable, we can still back out the frictionless response $E\left[\triangle z_e^*\right]$.
- Might still have implications for structural elasticities.

Optimization frictions

- To make out-of-sample predictions - *again*, need a "structural" model of $a(.)$.
- Model also needed if discontinuities do not create dominated regions e.g. kinks.
- Moreover, need additional moments to identify ϕ
- One approach - variation in size of kinks orthogonal to ϕ and e
  - e.g. differently sized kinks at different exogenously determined zs implies

$$B_1 = B_1(e, \phi, \mathsf{x}_1)$$
$$B_2 = B_2(e, \phi, \mathsf{x}_1)$$

Reference points

- Creation of a statutory threshold might lead to a focal point in certain choice problems.
- Implication: More bunching at thresholds than implied by structural e.
- Concern here is more about misinterpreting, rather than matching, data.
- Identification: There might be bunching at other earnings levels (such as round numbers in tax data for self employed).
    - Evidence of reference point effects, independent of statutory threshold.
    - Can be used to separate out the labor supply response at threshold.

Extensive margin

- Notches introduce discrete shifts in budget constraint, might cause people to drop out of work.
- The counterfactual density $\hat{h}_0$ at z^* will have to be purged of extensive margin response, to tease out the intensive margin response.
- Good news: In most models with participation constraints, small changes in $\triangle z^*$ do not lead extensive margin response.
- Useful implication for estimation: Excess mass (bunching) = Missing mass (holes)

Estimation: choosing z^+ and z^-

- Visual inspection
- In case of notches, diffusness of hole makes visual inspection of upper bound difficult.
  - Estimate the regression alongwith z^+ using the constraint that excess mass = missing mass (no extensive margin)
- "Fully automatic procedure" (Diamond & Persson, 2017)
  - Impose log concavity on density h_0 as well as densities of bunches and holes - shape restriction (like p).
  - Fit observed data under the constraint that excess mass = missing mass.
  - Implemented using bernstein polynomials and constrained NLS.

Applications

- Except tax responses/evasion, pensions, welfare programs, fuel economy policies, mortgage interest rates, education etc.
- Diamond and Persson (2017) estimate the effect of test score manipulation by teachers (to just pass students who might have had a bad test day) on later life outcomes such as earnings.
- Quick overview:
  - Sweden has a system where admission to high school depends on passing a grade 9 exam, which is scored locally.
  - Teachers have some discretion in grading, and incentives for passing more students.
  - Passing thresholds are determined after the exam.
  - Data shows bunching of test scores at these thresholds.

Diamond and Persson (2017)

Want to estimate: Causal effect of being manipulated across threshold (LATE)

- Step 0: Estimate $\hat{h}_0$ as described above.
- Step 1: From the non-manipulated portion of the scores distribution, estimate relationship between scores and earnings.
- Step 2: Find counterfactual scores for the manipulated portion -
  - i. Find counterfactual earnings using Step 1 and difference from actual data to find ATE in manipulated region
  - ii. Find the proportion of scores that were actually manipulated (proportion of "compliers")
- Step 3: Divide i by ii to get the LATE.
- Estimate large effects on earnings.