Economics 2450A: Public Economics and Fiscal Policy I

Section 7: Estimating Taxable Income Elasticities

Michael Droste

Fall 2022

Outline

1. Mid-Semester Feedback
2. Motivation and Brief History
3. Bunching estimator: Saez (2010)
   - Derivation
   - Estimation
   - Intuition
   - Problems with Identification
4. Notching
5. Other Applications / Discussion

Mid-Semester Feedback

- Survey from last week is still up! https://forms.gle/abYGfYKyubWDunRL8
- Thanks if you have provided feedback! There were two very clear suggestions for improvement:
  1. Practice problems
  2. Respect 75 minute section times
- These are both super reasonable and absolutely something I'm going to try and target.
- Practice problems covering the first half of the course (broadly, income/commodity taxation) will be posted this week.

Historical Motivation

- Taxable income and labor supply elasticities have featured prominently in our theories of optimal taxes/transfers. How can we estimate them?

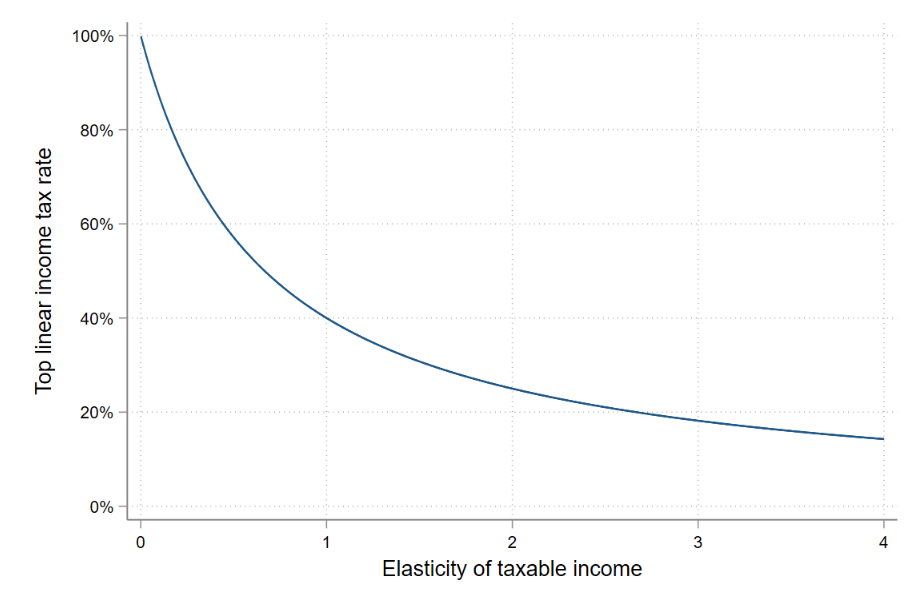
- Two main approaches in the 20th century:
  1. Structural approach (Burtless and Hausman (1978), Hausman (1981))
  2. Reduced-form approach (Lindsey (1987), Feldstein (1995))

- Structural approach: write down full model, with distributional and functional form assumptions (i.e. utility function, any relevant distributions over parameters, etc.), estimate with maximum likelihood given data (e.g. survey data on income and taxes).

- Reduced form approach: exploit tax reforms (treated as exogenous) in event study, DiD frameworks.

- Review question: What are potential drawbacks to each of the approaches above?

Optimal (Rawlsian/Revenue-Maxing) Top Linear Tax Rate vs. e



Bunching with Kinks: Friend or Foe?

- Basic labor supply model with piecewise linear tax schedule and household heterogeneity predicts that individuals ‘bunch’ at kinks in budget sets induced by, e.g., discrete changes in marginal tax rates. But little evidence of bunching in data (typically self-reported surveys) in the 1970-80s (prior to availability of administrative data).

- In structural work of 1970s-80s, theoretical prediction of bunching seen as a nuisance/artifact of survey data: explained away with measurement error terms.

- Saez (2010) turns this on its head with administrative data from IRS: points out there is a direct mapping from observed degree of bunching to taxable income elasticity in simple model. Can use kinked budget sets and observed degree of bunching to estimate elasticity using only cross-sectional variation (i.e. a single year of data).

Bunching with Kinks: Basic Tax Model (1/4)

- Individuals have utility over consumption and leisure. They have heterogeneous ability n (distributed with smooth density $f(n)$), earn pre-tax income $z = nl$, and face a budget constraint $c = z - T(z)$, where $T'(z) = t$ (for now).

- For simplicity, it is common to assume quasilinear isoelastic utility of the form:

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

- Maximizing utility subject to the linear budget constraint $c = (1 - t)z$ yields a first-order condition $z = n(1 - t)^e$, where e is the elasticity of taxable income with respect to the net of tax rate, i.e. $e = \frac{\partial z}{\partial(1-t)}\frac{(1-t)}{z}$.

- Let $h_0(z)$ denote the density of earnings under the constant marginal tax rate.

Bunching with Kinks: Basic Tax Model (2/4)

- Now suppose that the marginal tax rate increases to $t + \Delta t$ at z^*. This introduces a convex kink in individuals' budget frontiers, which we will visualize shortly.

- Individuals with $n \in [z^*/(1-t)^e, z^*/(1-t-\Delta t)^e]$ choose $z = z^*$, i.e. they bunch at the kink. The highest-ability person who bunches (the "marginal buncher") has ability $n = z^*/(1-t-\Delta t)^e$.

- Let $h(z)$ denote the density of earnings under the piecewise-linear tax schedule. Note $h(z) = h_0(z)$ for $z < z^*$. Denote $h(z)_-$, $h(z)_+$ the left and right limits (respectively) of earnings on either side of the threshold z^*.

Convex Kink and Bunching

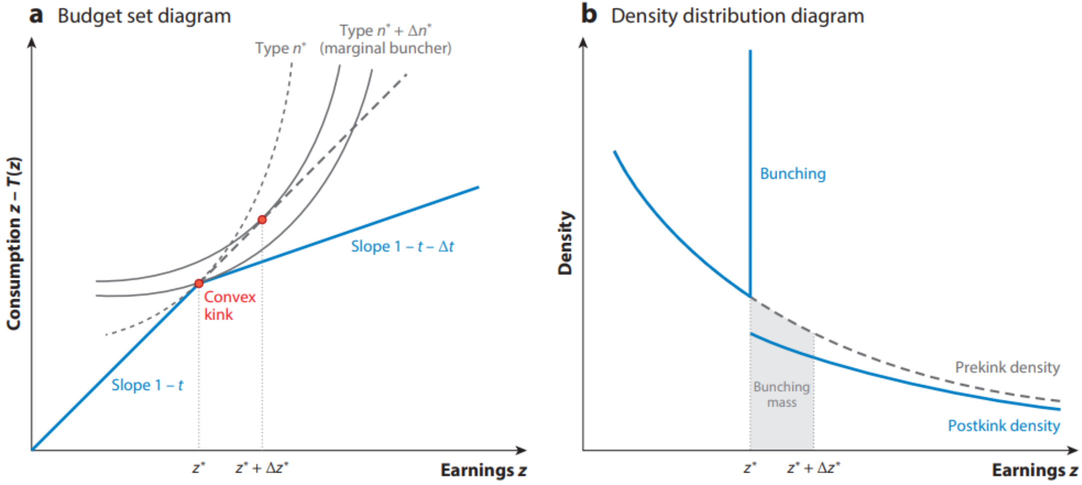


Figure 1

Bunching with Kinks: Basic Tax Model (3/4)

- Individuals who earn between z^* and $z^* + \Delta z^*$ under the linear tax schedule choose to bunch at z^* under the piecewise linear tax, with:

$$\frac{\Delta z^*}{z^*} = \left(\frac{1-t}{1-(t+\Delta t)}\right)^e - 1 \tag{1}$$

- The fraction of individuals bunching, which we denote B, can be expressed:

$$B = \int_{z^*}^{z^*+\Delta z^*} h_0(z)dz \approx \Delta z^* \frac{h_0(z^*) + h_0(z^* + \Delta z^*)}{2} \tag{2}$$

where we have employed a trapezoidal approximation to the integral in (2).

Bunching with Kinks: Basic Tax Model (4/4)

- Combining (1) and (2) leads to a key equation for bunching estimator:

$$B = z^* \left[\left(\frac{1-t}{1-(t+\Delta t)} \right)^e - 1 \right] \frac{h(z^*)_- + h(z^*)_+ / \left(\frac{1-t}{1-(t+\Delta t)} \right)^e}{2} \tag{3}$$

- Note that the tax parameters t, Δt and threshold z^* are directly observed.

- Thus, we only need to estimate the degree of bunching B and the density of income on either side of the threshold $h(z^*)_-$, $h(z^*)_+$ in order to use (3) to estimate e.

- We will return to estimating B in a second.

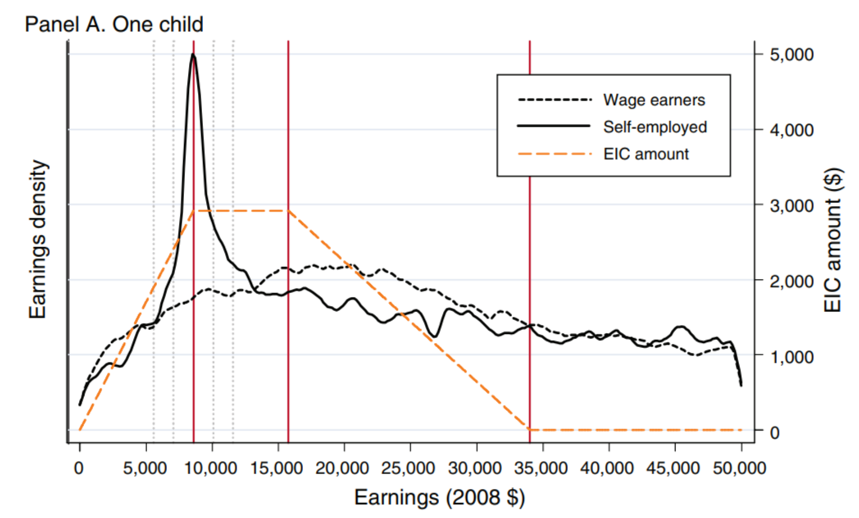
Bunching with Kinks: Some Comments

- We required a lot of structural assumptions (explicit and implicit) in order to derive (3), the key equation which allows us to estimate e from observed cross-sectional income data.

- Later, we'll discuss situations where identifying e becomes more complicated when we relax some of these assumptions.

- We only considered convex kinks, i.e. a piecewise linear tax schedule where the marginal tax rate *increases* at z^*. In the presence of non-convex kinks (i.e. the marginal tax rate decreases above z^*), this simple model predicts a "hole" rather than "bunching". This has no impact on equation (3), just the graphical interpretation.

Example: Observed Bunching in EITC Schedule



Estimation and Inference

- Estimation:
  1. Estimate densities on each side of threshold $\hat{h}(z^*)_-$, $\hat{h}(z^*)_+$ (easy)
  2. Estimate bunching parameter $\hat{B}$ (hard)
  3. Plug in estimates above into (3) and solve for estimate of elasticity, $\hat{e}$ (easy)

- Inference (standard errors, hypothesis tests involving $\hat{e}$) is straightforward: can use a bootstrap (re-estimating densities and bunching parameter many times to compute many estimates, approximating sampling distribution) or delta method.

- Key challenge is to estimate B: implies a choice of counterfactual density/

Estimating the Counterfactual Density

- Two basic approaches have been used to estimate *B*.

- Simplest approach: Uniform counterfactual density
  - Specify exogenous bandwidth δ around the kink, assume uniform density in band $[z^* - \delta, z^* + \delta]$.
  - Reasonable if considering a "small" kink - maybe not otherwise (Saez 2010)

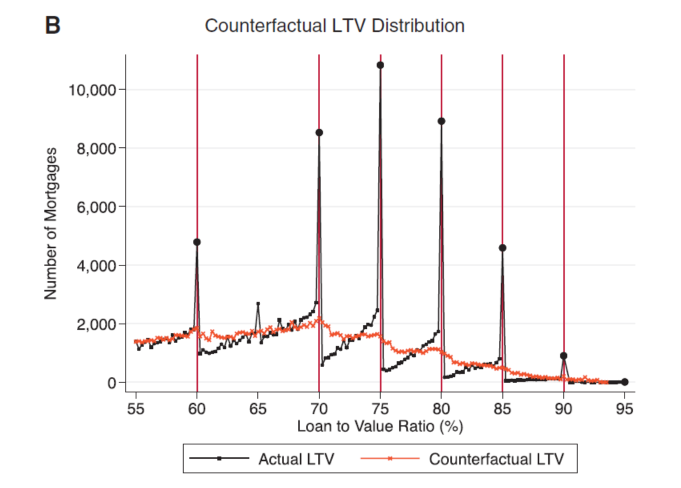
- More sophisticated approach: Flexible polynomial fit
  - Also relies on bandwidth choice δ and order of polynomial *p*
  - Take advantage of information about how distribution is changing *outside* interval $[z^* - \delta, z^* + \delta]$ to inform counterfactual density inside that interval.
  - Described more fully by Chetty et al. (2011)

Counterfactual Density: Alternatives

- Use panel data of individual choices over time to get individual-level counterfactual
- Use empirical density absent the policy
  - Before/after a policy change
  - OR across agents subject to/not subject to the policy
- Using this as a counterfactual density eats up an empirical moment that could be used to identify extra parameters (ϕ,r, etc.)
  - $\Rightarrow$ potential to link all these moments together in GMM?

Counterfactual Density: Example

- Best et al 2020 use Loan-to-Value Ratio *before the agent (mortgage refinancer) acts* (with an adjustment for their context)



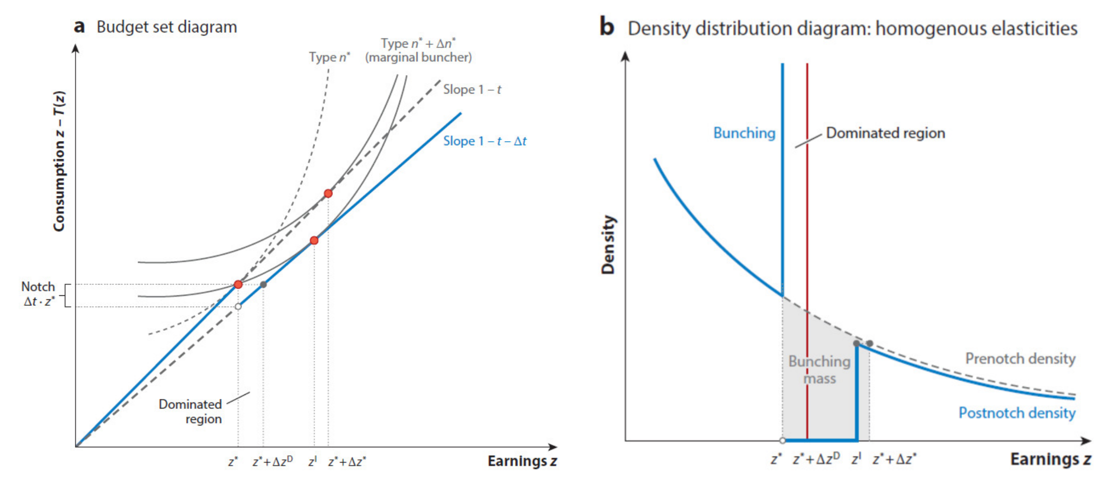
Notching Estimators

- Sometimes policies induce discontinuities rather than kinks in budget sets. It turns out that these discontinuities can be used to set up a similar design, sometimes called a notching estimator.

- Convex notches induce dominated regions. For instance, if the average tax rate increases discontinuously at a threshold z^*, there is some neighborhood $[z^*, z^* + \Delta z^*]$ where individuals can increase both consumption and leisure by moving to the threshold z^*.

- The same method of deriving the “key bunching equation” can be followed to derive a distinct expression mapping excess density around a notch threshold to the elasticity of interest.

Notch (upward) (Kleven 2016 Figure 2)



Identification Problems

- e is point-identified under the structural model we specified.
- However, we relied on a number of implicit structural assumptions that are quite strong; when these are relaxed, the bunching estimator generally does not (point) identify the parameter of interest unless additional empirical moments are introduced and/or additional structural assumptions are imposed.
- We will discuss three important extensions to the basic model that have received some attention in the literature: optimization frictions, reference dependence, and preference heterogeneity.

Identification Problems: Optimization Frictions

- Optimization frictions arise when individuals are not on the first-order condition characterizing optimal behavior; for instance, an individuals labor supply decision may be subject to costly adjustment, search frictions, or inattention.

- Empirical support for the relevance of optimization frictions:
  - Bunching is larger when the kink/notch is larger, more salient, or stable over time
  - When it exists, bunching is relatively diffuse: density doesn't spike at thresholds
  - Sometimes observe people in dominated regions induced by notches

- Problems introduced by optimization frictions:
  - Attenuates bunching (potentially not a problem with bandwidth chosen correctly)
  - Gap between "observed" and structural elasticities
  - Failure of order condition for identification: One empirical moment (degree of bunching B) and two structural parameters (elasticity e and the optimization friction parameter ϕ). Potentially many combinations of (e, ϕ) consistent with observed B.

Identification Problems: Optimization Frictions (continued)

- Obtain additional data moments that depend on (e, ϕ)
- For **upward notches**: use the dominated region, which should be empty in a frictionless world, to estimate the share of nonoptimizers $a^*(\phi)$
- For **downward notches**: (more parametrically) rule out extreme preferences
- For **kinks**: use any variation in the size of the kink orthogonal to (e, ϕ), assuming the same elasticity e and friction ϕ throughout
  - Multiple kinks of varying sizes (i.e. federal income tax schedule)
  - Changes in the size of a kink over time (i.e. tax reforms)

Identification Problems: Preference Heterogeneity

- We have implicitly assumed homogeneous preferences in a region around the kink (uniform counterfactual density).

- Newey et al. (2019): in canonical bunching design, e not identified when preference heterogeneity is unrestricted. Intuitively: with heterogeneous preferences, bunching could be either behavioral responses or preference for income around kink point.

- This is an important critique, because it cuts against a key virtue of the bunching approach, which is that it doesn't require variation in budget sets within individual.

- Alternatives: can use policy changes, differential exposure to policy to estimate counterfactual treatment. At the very least: always need to be clear about counterfactual density!

Identification Problems: Reference Points

- Notches and kinks may also serve as reference points (prospect theory)
- Potential explanation for no empirical observation of holes?
- Could *amplify* bunching $\Rightarrow$ estimated elasticity *overstates* structural elasticity
- e.g. round-number bunching of reported taxable income, housing prices
  - $B(e, \phi, r, x)$, where r is the reference point effect
  - Better estimation of e, ϕ means estimating a data moment of r: e.g. estimating bunching at round numbers that are not also notches or kinks
  - In practice: use round-number fixed effects

Applications of Bunching

- We have described bunching in the “canonical” context of estimating taxable income elasticities, as in Saez (2010). In the last decade, researchers have used bunching estimators to study behavioral responses in a variety of settings outside of the tax literature.

- A few examples: Pensions, social insurance, welfare programs, labor regulation, minimum wages, education, electricity prices, fuel economy policy, cellular service prices, mortgage interest rates

- See Kleven (2016) for other interesting examples

Wrapping Up

- What have we learned? Bunching estimators are an interesting tool for estimating structural parameters / behavioral responses... But this approach requires strong structural assumptions!

- Food for thought:
  - How much do optimization frictions matter in practice? What is the bias from ignoring them? Do policymakers care about the ’true’ structural elasticity or the observed elasticity?
  - Are kinks/notches really exogenous?
  - If policies are chosen endogenously based on agents’ “demand” for kinks and notches, then (1) how do we adjust for this? (2) in which settings is this more likely to be a problem?

References

- Best, Cloyne, Ilzetzki, & Kleven (2020). Estimating the Elasticity of Intertemporal Substitution Using Mortgage Notches. The Review of Economic Studies, 87(2), 656–690.
- Blomquist, Newey, Kumar, & Liang (2017). On Bunching and Identification of the Taxable Income Elasticity. NBER Working Paper #24136.
- Chetty, Friedman, Olsen, & Pistaferri (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.
- Chetty, Friedman, & Saez (2013). Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings. American Economic Review, 103(7), 2683–2721.
- Einav, Finkelstein, & Schrimpf (2017). Bunching at the kink: Implications for spending responses to health insurance contracts. Journal of Public Economics, 146, 27–40.
- Kleven (2016). Bunching. Annual Review of Economics, 8(1), 435–464. *(main source for this presentation)*
- Kleven, & Waseem (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. The Quarterly Journal of Economics, 128(2), 669–723.
- Saez (2010). Do Taxpayers Bunch at Kink Points? American Economic Journal: Economic Policy, 2(3), 180–212.