Optimal deductibility:

Evidence from a bunching decomposition

Steven Hamilton

The George Washington University

steven_hamilton@gwu.edu

steven-hamilton.com

@SHamiltonian

Overview

To what degree should taxpayers be allowed to claim deductions?

Overview

To what degree should taxpayers be allowed to claim deductions?

- *What parameter do I need to identify?*
  - New way of thinking about deduction policy choice.
  - Under separability, deduction elasticity is key parameter.

Overview

To what degree should taxpayers be allowed to claim deductions?

- *What parameter do I need to identify?*
  - New way of thinking about deduction policy choice.
  - Under separability, deduction elasticity is key parameter.
- *How can I identify the deduction elasticity?*
  - New method to decompose bunching in taxable income.
  - Exploit year-on-year change in bunching incentives.

Overview

To what degree should taxpayers be allowed to claim deductions?

- *What parameter do I need to identify?*
  - New way of thinking about deduction policy choice.
  - Under separability, deduction elasticity is key parameter.
- *How can I identify the deduction elasticity?*
  - New method to decompose bunching in taxable income.
  - Exploit year-on-year change in bunching incentives.
- *What is the deduction elasticity in practice?*
  - New admin. data to exploit Australian policy discontinuity.
  - Deduction elasticity -0.45 and gross-income elasticity 0.04.

Literature

- Elasticity of taxable income (ETI)
  - Feldstein (1995, 1999); Slemrod (1998); Chetty (2009); Doerrenberg, et al. (2015)
  - *I show the ETI is not sufficient for a different tax instrument.*
- Bunching
  - Saez (2010); Chetty, et al. (2011); Kleven and Waseem (2013)
  - *I develop the first method to decompose bunching.*
- Deductions
  - Doerrenberg, et al. (2015); Schächtele (2016); Paetzold (2017)
  - Best, et al. (2015); Bachas and Soto (2016)
  - *I provide a new estimate of the deduction elasticity.*

A model of optimal deductibility

Taxpayer's problem

Taxpayer chooses consumption, c, gross income, y, and deductions, d, given tax rate, τ, and deductibility rate, ρ, to maximise utility. Taxable income is $z = y - \rho d$.

$$\max_{c,y,d} \; u(c, y, d) \quad \text{s.t.} \; c \leqslant y - d - \tau \cdot (y - \rho d),$$

which yields:

$$\frac{u_y}{u_d} = -\frac{1-\tau}{1-\rho\tau}.$$

Government's problem

The government chooses τ and ρ to maximise indirect utility, $v(\tau, \rho)$, and the external value of deductions, $\Phi(d)$.

$$\max_{\tau,\rho} \; v(\tau,\rho) + \Phi(d(\tau,\rho)) \quad \text{s.t.} \; \tau \cdot (y - \rho d) \geqslant R.$$

Identifying welfare impact of ρ requires variation in ρ:

$$\underbrace{\frac{\partial R}{\partial \rho} + \tau d}_{\text{Revenue leakage}} \quad = \tau \cdot \left(\frac{\partial y}{\partial \rho} - \rho \cdot \frac{\partial d}{\partial \rho} \right),$$

but we don't commonly observe such variability.

Quasilinear, isoelastic & separable utility

- More common to observe variation in τ.
- But under quasilinear, isoelastic, and separable utility, variation in τ and in ρ have the same effect on deductions:

$$\begin{aligned} u(y,d) = {} & y - d - \tau \cdot (y - \rho d) \\ & - \frac{n_y}{1+1/e_y} \cdot \left(\frac{y}{n_y}\right)^{1+1/e_y} + \frac{n_d}{1+1/e_d} \cdot \left(\frac{d}{n_d}\right)^{1+1/e_d}. \end{aligned}$$

Optimal deductibility rate

With this functional form, the optimal deductibility rate is:

$$\rho^*(\tau) = \frac{1}{\tau} \cdot \frac{1 - \lambda^g - \Phi'(d) \cdot e_d}{1 - \lambda^g - \lambda^g \cdot e_d},$$

which has the Ramsey (1927) inverse-elasticity form.

A bunching decomposition method

Standard bunching method

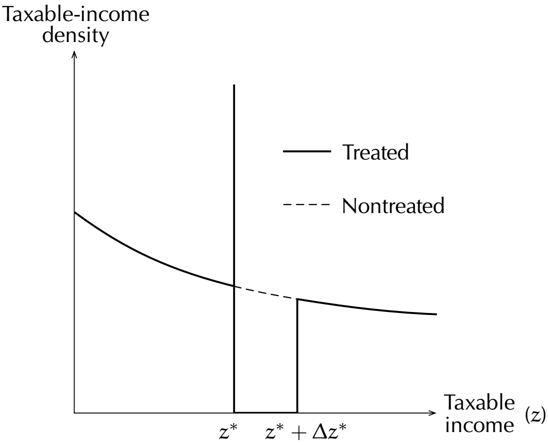


Figure: Densities with and without the notch

Bunching decomposition method

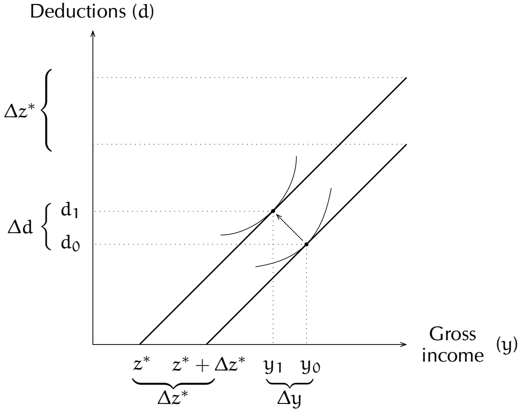


Figure: Changes in deductions and gross income due to the notch

Deriving the item elasticities

For an optimising buncher:

- in the absence of the notch:

$$y_0 = n_y(1-t)^{e_y} \qquad\qquad d_0 = n_d(1-t)^{e_d}$$

- and in the presence of the notch:

$$\left(\frac{y_1}{n_y}\right)^{1/e_y} = \left(\frac{d_1}{n_d}\right)^{1/e_d},$$

such that $y_1 - d_1 = z^*$.

Deriving the item elasticities

These first-order conditions yield:

$$\frac{e_y}{e_d} = \frac{\ln y_1 - \ln y_0}{\ln d_1 - \ln d_0}.$$

Note the elasticity of taxable income is the weighted average:

$$e = \frac{y}{z} \cdot e_y - \frac{d}{z} \cdot e_d$$

Combining these yields:

$$e_d = e \cdot \frac{\ln\left(\frac{\Delta d}{d_0} + 1\right)}{\frac{y_0}{z_0} \cdot \ln\left(\frac{\Delta y}{y_0} + 1\right) - \frac{d_0}{z_0} \cdot \ln\left(\frac{\Delta d}{d_0} + 1\right)} \approx e \cdot \frac{\Delta d}{\Delta z} \cdot \frac{z}{d}.$$

Deriving the item elasticities

- Need to estimate:

$$\hat{e}_d = \hat{e} \cdot \frac{\hat{\mathbb{E}}[\, d_1 - d_0]}{\hat{\mathbb{E}}[\, z_1 - z_0]} \cdot \frac{\hat{\mathbb{E}}[\, z_0]}{\hat{\mathbb{E}}[\, d_0]},$$

  all conditional on $z^* \leqslant z_0 \leqslant z^* + \Delta z^*$.

- ETI, ATEs, and average outcomes under nontreatment.
- Need to observe treatment and comparison groups.

Empirical analysis

Institutional settings

Medicare Levy Surcharge

- 1% tax on childless singles without private health insurance, and with taxable income above AU$50,000.
- Different threshold for couples, based on joint income.
- In 2009, threshold was increased to $70,000.

Institutional settings

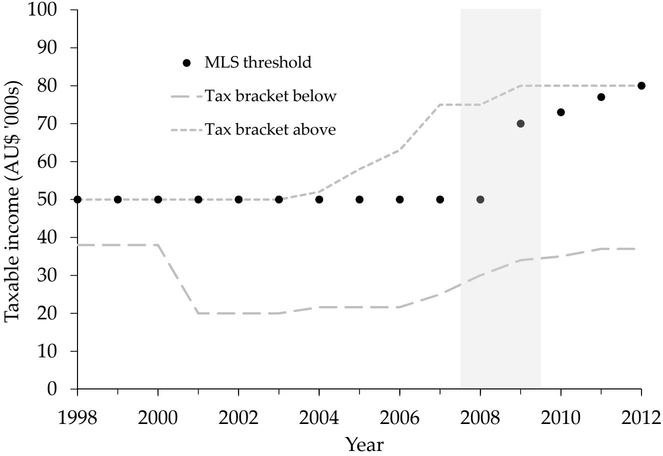


Figure: MLS threshold over time

Data

- Use Australian Treasury administrative tax data.
- 16% sample (2 million observations in total).
- Exclude married people, those with children, and those covered by health insurance.
- Within $2,250 income range considered and conditional on characteristics, dataset contains 80,000 observations.

Identification strategy

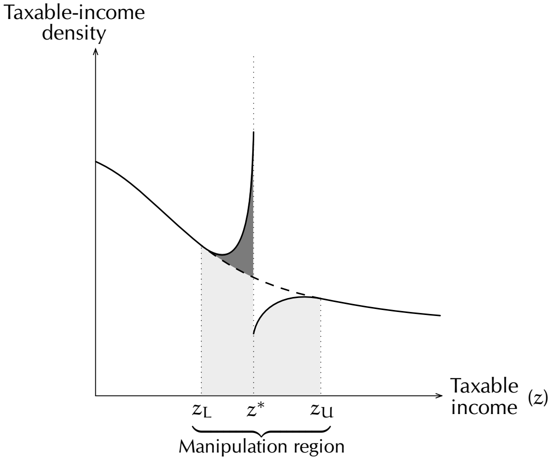


Figure: Bunching in the manipulation region.

Identification strategy

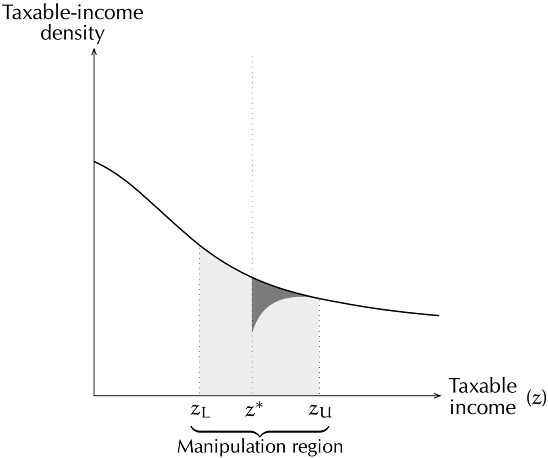
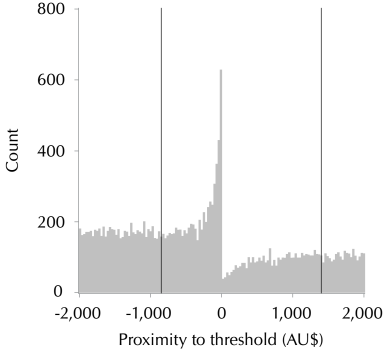
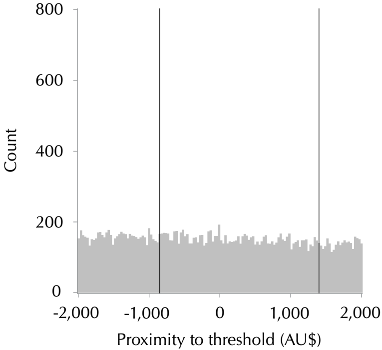


Figure: Bunching in the manipulation region.

Determining manipulation region



(a) 2008 histogram



(b) 2009 histogram

Determining manipulation region

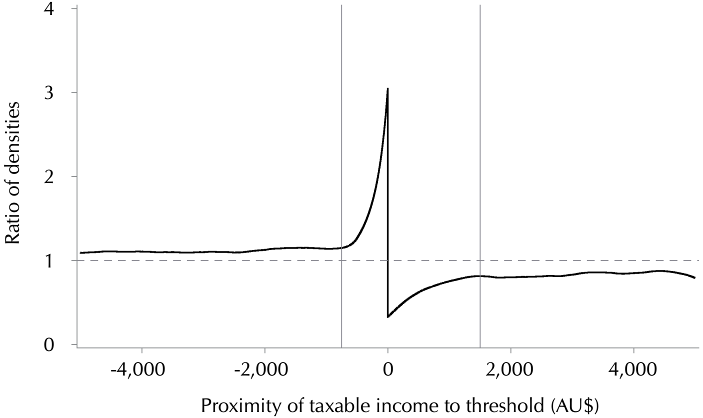


Figure: Density ratio from local-logit-regression predicted values

Determining placebo region

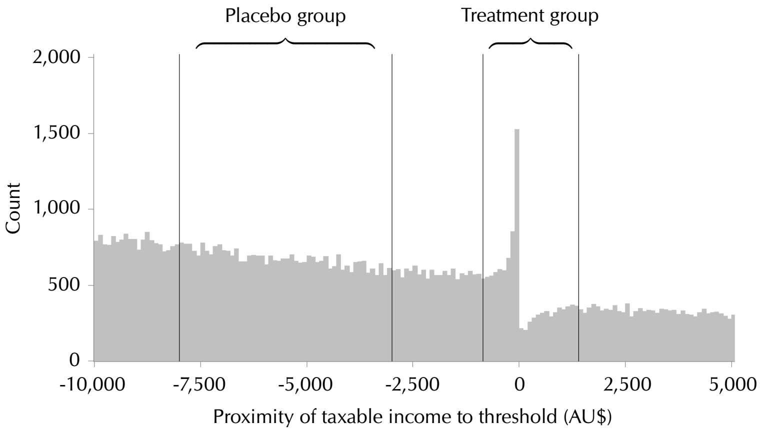


Figure: Density under treatment in treatment and placebo regions

Determining placebo region

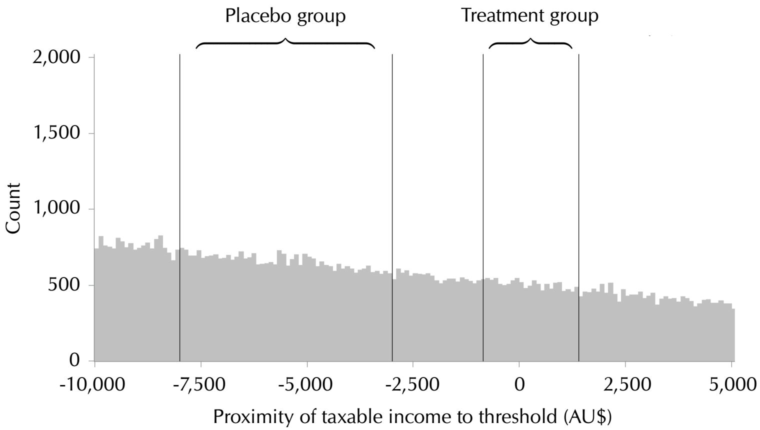
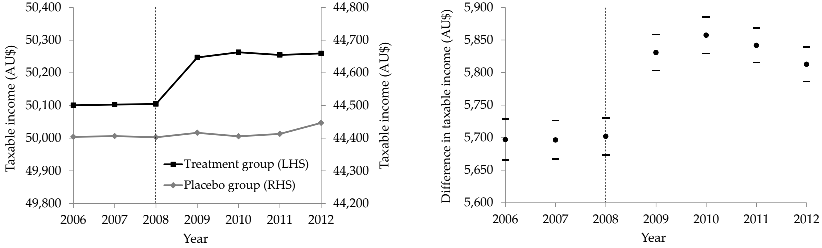


Figure: Density under nontreatment in treatment and placebo regions

Estimating ATEs



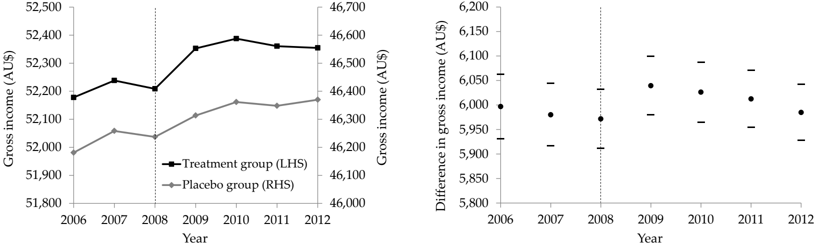
(a) Levels of taxable income

(b) Difference in taxable income

Pretrend-corrected DiD estmates:

- $127.10 (26.61) among all taxpayers.
- $526.73 (79.75) among bunchers only.

Estimating ATEs

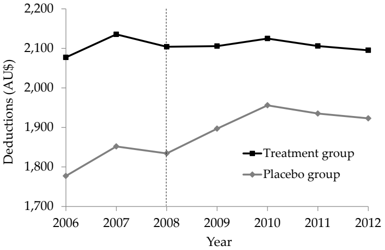


(a) Levels of gross income (b) Difference in gross income

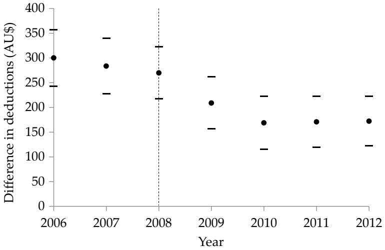
Pretrend-corrected DiD estmates:

- $82.05 (56.74) among all taxpayers.
- $340.03 (230.60) among bunchers only.

Estimating ATEs



(a) Levels of deductions



(b) Difference in deductions

Pretrend-corrected DiD estmates:

- −\$45.05 (49.74) among all taxpayers.
- −\$186.70 (215.55) among bunchers only.

Estimating ATEs

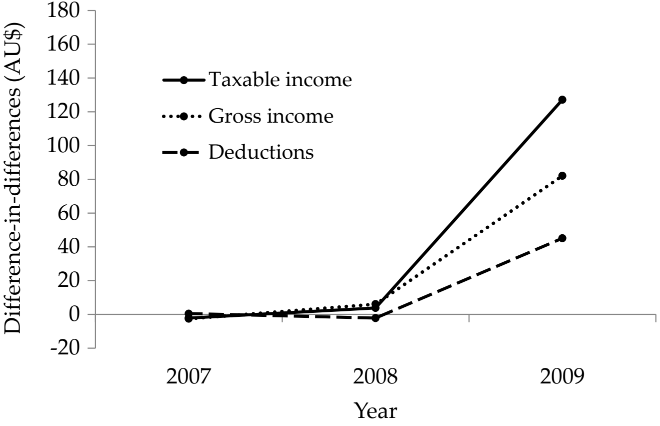


Figure: Pretrend-corrected differences-in-differences over time.

Bottom line

	% of TI	% of ΔTI	Item elasticity w.r.t.	
			Taxable income	Net-of-tax rate
Gross income	104.71	64.55	0.62	0.04
Deductions	4.71	35.45	−7.53	−0.45

Table: Estimated item elasticities (ETI is 0.06).

- With 20% efficiency loss, marginal dollar of deductions requires 68¢ in external benefits for optimal full deductibility.
- If external benefits were 30¢, then $\rho^* = 0.34$.

Conclusion

- Under separability, tax-rate variation proxies for deductibility-rate variation.
- Decompose ETI via relative proportional changes of items and taxable income in bunching.
- Deductions account only for 5% of taxable income, but 35% of the response of taxable income to taxes.
- Deduction elasticity -0.45 and gross-income elasticity 0.04.
- Because deductions are granted at a high welfare cost, lowering deductibility is likely to raise welfare.

Fin