Optimal deductibility: Theory, and evidence from a bunching decomposition

Steven Hamilton

George Washington University

steven_hamilton@me.com

steven-hamilton.com

July 13, 2018

Overview

To what degree should taxpayers be able to claim deductions?

- *What are the necessary statistics?*
  - Ramsey-style model of optimal deductibility.
- *How can these be identified?*
  - Bunching decomposition method.
- *What do the data tell us?*
  - Application to change in threshold of an Australian notch.

A model of optimal deductibility

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

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

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

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

Optimal deductibility rate

Assuming quasilinear, isoelastic, and separable utility:

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

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 usual Ramsey (1927) inverse-elasticity form.

Standard bunching method

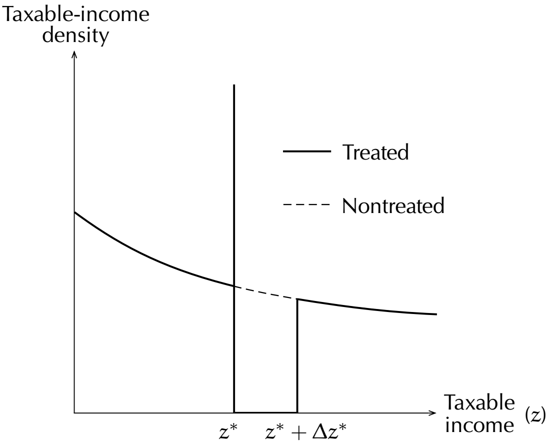


Figure: Densities with and without the notch

Item choices when bunching

- Bunching decision is discrete, by comparing utilities.
- Bunchers reduce taxable income, equating marginal utilities of gross income and deductions.
- Use these conditions to solve for the deduction elasticity:

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

where the deduction elasticity is relative proportional change of deductions and taxable income, scaled by ETI.

Deriving the item elasticities

- Need to estimate five objects:

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

  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.

Institutional settings

- 16% sample of ATO unit-record data.
- Largest deductible expenses: work-related car, clothing, travel, and education; charity; tax preparation.
- Medicare Levy Surcharge
  - 1% notch for childless singles w/o health insurance, and with taxable income above $50,000.
  - In 2009, threshold increased to $70,000.

Institutional settings

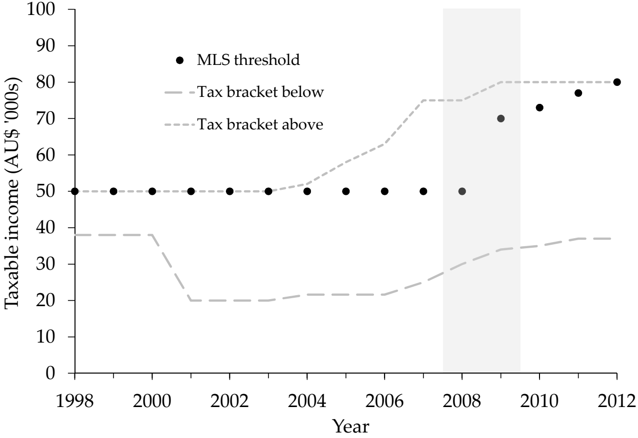


Figure: MLS threshold over time

Identification strategy

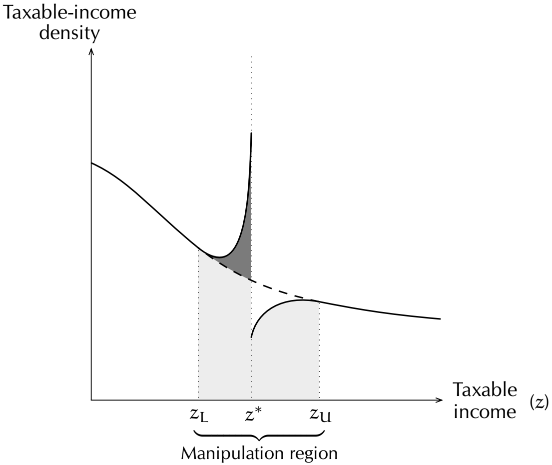


Figure: Bunching in the manipulation region.

Identification strategy

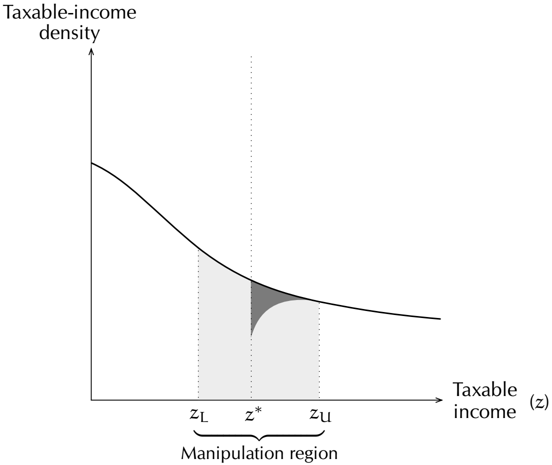
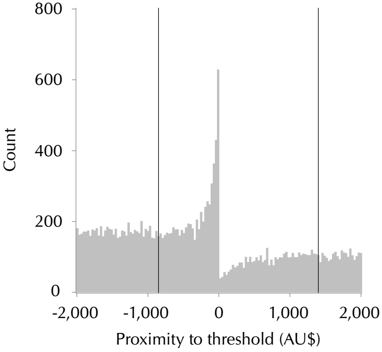
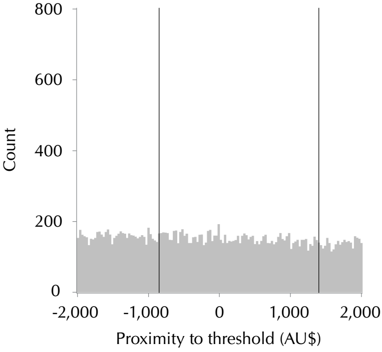


Figure: Bunching in the manipulation region.

Determining manipulation region



(a) 2008 histogram



(b) 2009 histogram

Determining manipulation region

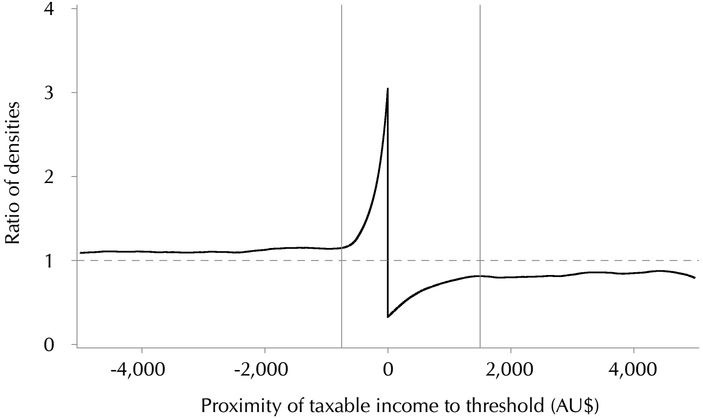


Figure: Density ratio from local-logit-regression predicted values

Determining placebo region

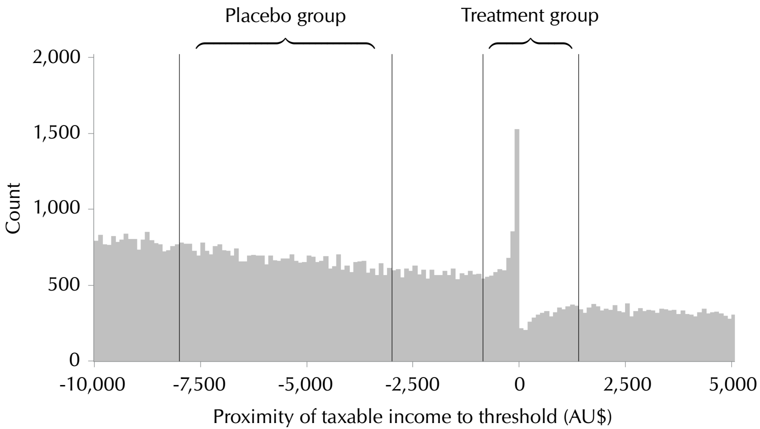


Figure: Density under treatment in treatment and placebo regions

Determining placebo region

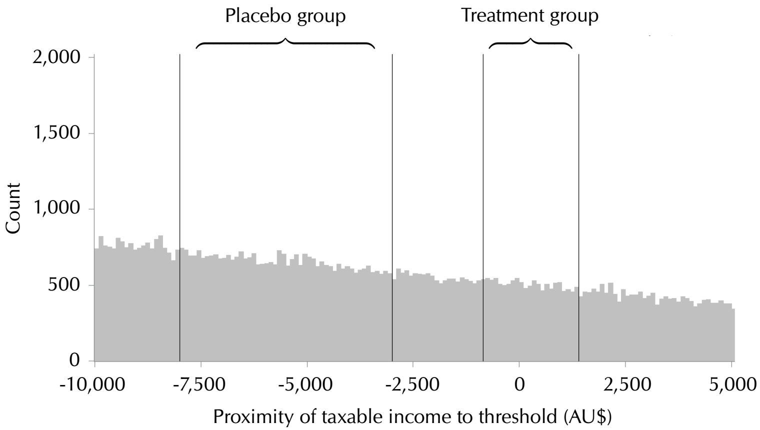
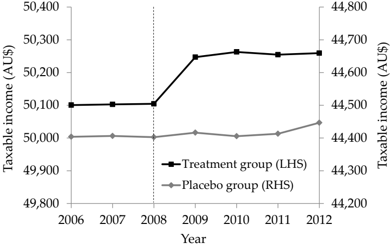
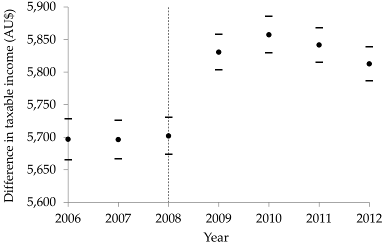


Figure: Density under nontreatment in treatment and placebo regions

Estimating ATEs



(a) Levels of taxable income

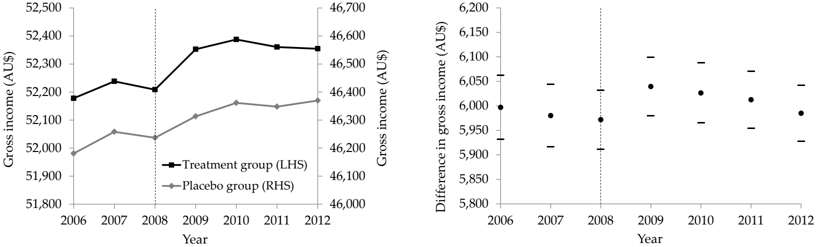


(b) Difference in taxable income

Pretrend-corrected DiD estmates:

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

Estimating ATEs

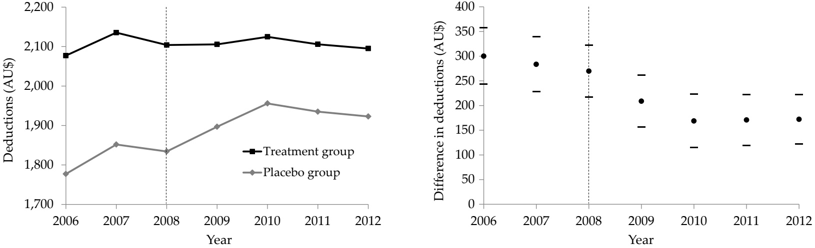


(a) Levels of gross income (b) Difference in gross income

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

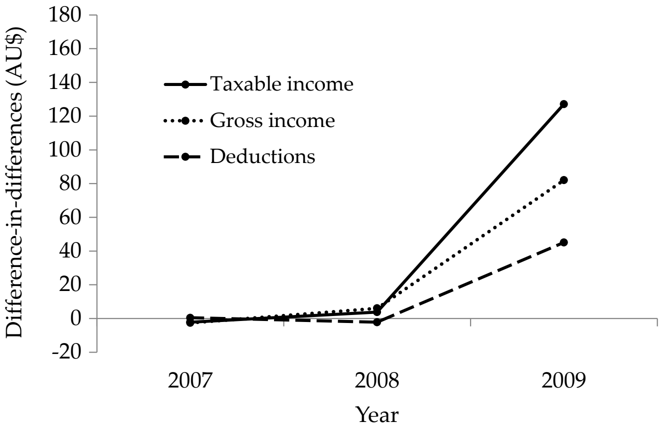


Figure: Pretrend-corrected differences-in-differences over time.

Sufficient statistics

	% 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

- ‘Sufficient statistics’ approach: use structural model to derive parameters for optimal policy, identified via policy 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.
- Deductions are granted at a high welfare cost.

Fin

Steven Hamilton

George Washington University

steven_hamilton@me.com

steven-hamilton.com