Ben Lockwood and Alex Rees-Jones
Wharton

Problem Set 2

Spring 2025

Due by class time on Wednesday, April 16. Please submit electronically, via emailed PDF to ben.lockwood@wharton.upenn.edu and alre@wharton.upenn.edu. You may collaborate with other students; if you do, please list your collaborators in your submitted problem set. Each student should send a separate submission.

Part 1: Elasticity Estimation

In this exercise, we'll make sure you have a sense for how to conduct basic bunching-based elasticity estimation.

To begin, examine the file middle_kink.csv. It contains data from the paper "DiffuseBunching with Frictions: Theory and Estimation" by Anagol, Davids, Lockwood, and Ramadorai. You should download the paper (available on Ben's website) for additional context. The file reports the number of small businesses reporting incomes in each bin around the "middle kink" tax bracket threshold discussed in that paper.

As in problem set 1, you should submit answers to the problems below, as well as replication-ready code (in Stata or R) that produces those results.

In answering the following problems, there are a number of degrees of freedom that we have left unspecified, such as the polynomial order for the counterfactual distribution in part (c). That ambiguity is intentional—make what you view as reasonable assumptions, as you would if you were conducting this research from scratch.

a) Make a publication-quality histogram of the earnings distribution. This means that all axes should be nicely and clearly labeled and the whole thing should look good. Draw a vertical line at the point with a kink in the tax schedule.

b) Saez (AEJ:Policy 2010) provides a technique for inferring elasticity from excess mass around a kink point. Use this strategy to generate an estimate of the elasticity in this dataset. Your computation should be based on equation (5) in Saez (2010). It is fine if you generate estimates of items (1)-(4) that he mentions just after equation (5) and then use an equation-solver to generate your estimate of *e.*

   Your "deliverable" for this problem is to write out your version of equation 5, with your empirical numbers plugged in, for the data in middle_kink.csv, and to report the *e* that solves that equation. You should explain how you are generating your numbers for B, $h(z^*)_-$, and $h(z^*)_+$.

c) When researchers have to estimate excess mass, they commonly use the approach proposed by Chetty, Friedman, Olsen, and Pistaferri (QJE 2011). These authors provided a Stata package that executes their approach. It can be downloaded on Raj Chetty's website: click on this paper, and then "Bunching Estimation Code." Alternatively, you can try using the bunchr package in R.

   Use this code to generate estimates of the excess mass around the kink point in this data. Your "deliverable" is a figure displaying the raw data and the counterfactual density, as well as an estimate of the elasticity from this package.

d) Does the elasticity you estimated in part (b) match the one produced in part (c)? Why or why not? How do these estimates compare to the preferred estimate in the working paper by Anagol et al.? Why might they be different?

Part 2: Essay Response [A few paragraphs of prose expected]

Should we tax capital? Why or why not? Your answer should engage with the theoretical debate on this question.