A Big Data Approach to Optimal Taxation

Richard W. Evans
Brigham Young University

July 2014

Current Interesting Tax Stuff

- Baker, Bejarano, Evans, Judd, and Phillips (2014)
  - "A Big Data Approach to Optimal Sales Taxation," NBER Working Paper, #20130 (May 2014).
- BYU Macroeconomics and Computational Laboratory (MCL)
  - Undergrad "boot camp"
  - mentored research
  - conference hosting/conference presentations
- AEI Open Source Policy Center
  - Provide open source option for static and dynamic scoring models

Big Data Research: Definition and Others

Definition: Big Data

Any repository of data that is either large enough or complex enough that distributed and parallel input and output approaches must be used.

- Einav and Levin (2013), new opportunities for big data in economics
  - Primarily empirical, searching for patterns among large datasets

Big Data Research: Our approach

- We use big data techniques as solution method to theoretical models
- Ideal for:
  - High dimensional heterogeneity
  - Nonconvex optimization
  - Model and objective uncertainty

Many different economic tax models

- Tax model universe broad, varied, disparate
- No grand unifying theory
- Underlying macroeconomic systems too complex
- Include relevant structures to question
- Exclude less salient structures to question

Human body analogy

- Skeleton and crash-test dummy are two approximations of human body.
- Different models for different uses

Dimensions of tax model complexity

1. Partial equilibrium vs. general equilibrium
2. Static vs. dynamic
3. Deterministic vs. stochastic
4. Reduced form vs. structural
5. Linear vs. nonlinear
6. Positive vs. normative
7. Low heterogeneity vs. high heterogeneity

No model has all seven dimensions

- Cannot solve all seven analytically or computationally
- All models trade off some dimensions of complexity and richness for others

Sampling of tax models

- New Dynamic Public Finance DSGE
- New Keynesian DSGE
- Big input-output models
- CBOLT
- JCT microsimulation static scoring

Big Data

- Big data approach focuses on:
  - structural models
  - nonlinear solution methods
  - normative welfare analysis
  - high dimensional heterogeneity
  - nonconvex optimization problems
- Big data approach goes light on (for now):
  - partial equilibrium
  - static
  - deterministic

Relaxing Strong Assumptions

- “All” structural tax models make strong assumptions to induce convexity

Deaton (1977, p. 310)

“The result rests on strong simplifying assumptions in order to avoid the complexity of the general case; in particular, consumer bahaviour has been restricted by the use of linear Engel curves and by permitting only very limited substitution between commodities. Relaxing either of these could alter quite fundamentally the nature of the empirical results.”

- We are relaxing those assumptions to see if results hold

Steps to our approach

1. Choose dimensions of household/firm heterogeneity θ_i
2. Households/firms make choices to maximize individual objective $\boldsymbol{x}_i$
3. Given tax schedule $\boldsymbol{\tau}_j$:
   - households/firms make optimal choices $\boldsymbol{x}(\theta_i, \boldsymbol{\tau}_j)$
   - utility/profits at those optimal choices $u(\theta_i, \boldsymbol{\tau}_j)$
   - household/firm taxes paid are determined $r(\theta_i, \boldsymbol{\tau}_j)$

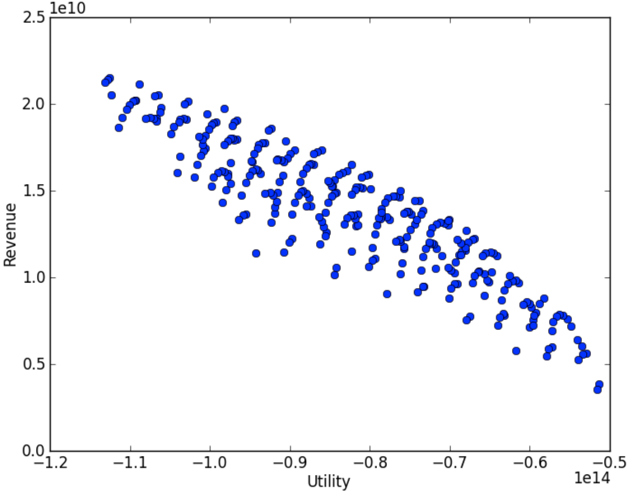
Hard Part: storing the grid of responses

- Must choose grid over both household/firm types $\boldsymbol{\theta}$ and over possible tax schedules
- Example $\boldsymbol{\theta} = \{\text{ability}, \text{labor dislike}\}$, 10 types each
- Assume two goods, each with tax $\tau_1, \tau_2 = 0\%, 10\%, ...90\%$
- **Implies** $10,000 = 100 \times 100$ **optimization problems**

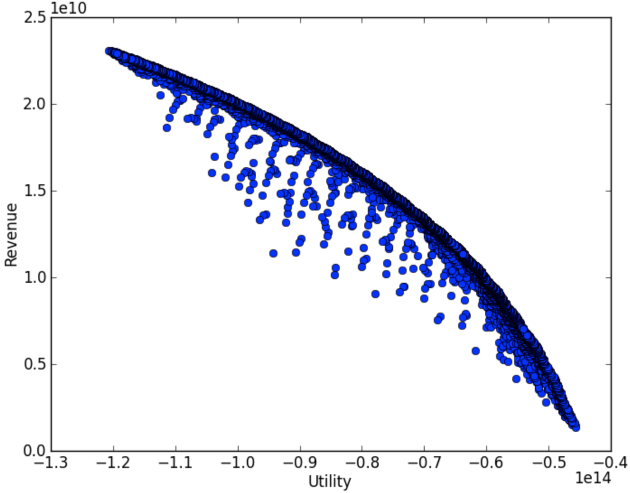
Steps to our approach

4. After grid of optimal behavior is stored, choose population distribution $\Gamma(\boldsymbol{\theta})$
5. Add up total utility for each tax policy $U_j = \sum_i u(\theta_i, \boldsymbol{\tau}_j)$
6. Add up total tax revenue for each tax policy $R_j = \sum_i r(\theta_i, \boldsymbol{\tau}_j)$
7. Plot all the (U_j, R_j) combinations for each tax policy
8. Sweep out strictly dominated policies
9. Refine search around policies on the frontier

Total utility-total revenue (U_j, R_j) combinations



Total utility-total revenue (U_j, R_j) frontier



What you get

- For given total revenue R_j, optimal tax policy τ_j that maximizes societal welfare U_j at that tax policy
- Parallel computing required to compute all the individual responses for all the points in the tax grid $\boldsymbol{\tau}$
- Big data techniques to store household/firm responses in database
- Big data techniques to calculate (U_j, R_j) for all $\boldsymbol{\tau}_j$

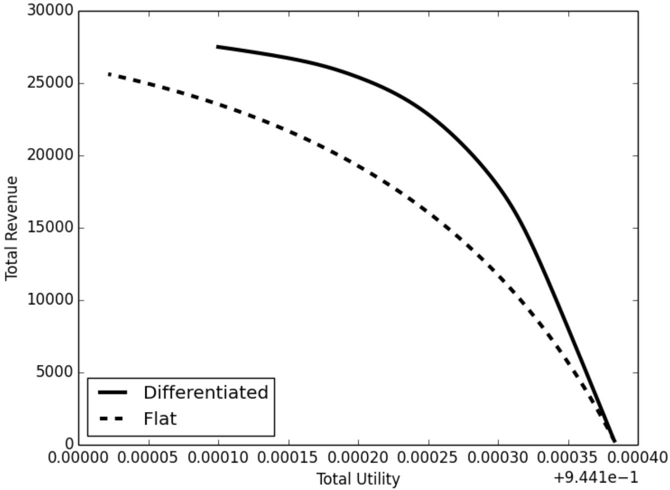
Sales Tax Example

Research questions

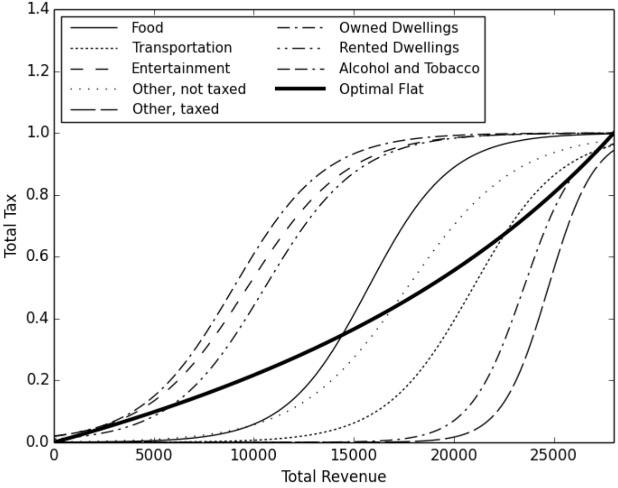
1. How big is the revenue difference between differentiated sales tax and flat tax?
2. How big is the loss exempting broad services category?

- Calibrate heterogeneity to:
  - U.S. household income distribution
  - Estimated U.S. price markups
  - U.S. consumption category shares by consumer income

Differentiated vs. flat tax frontiers



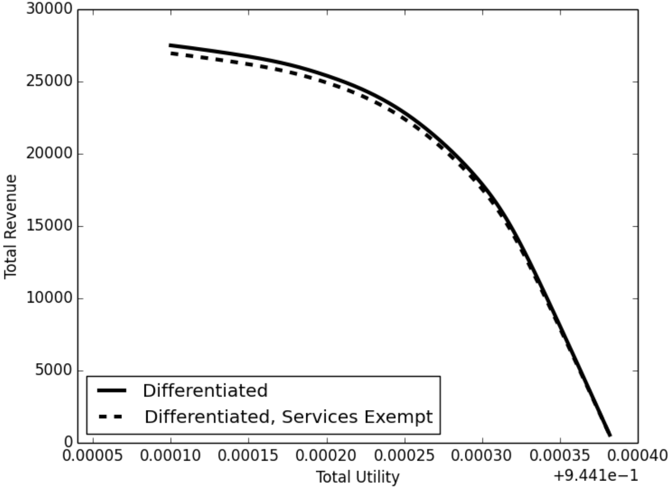
Differentiated vs. flat tax rates



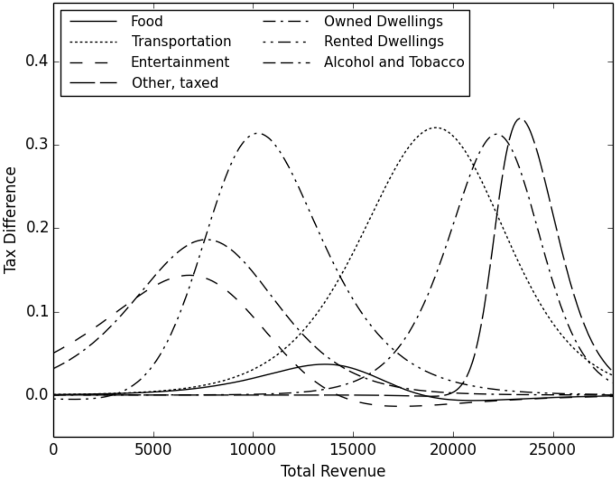
Differentiated vs. flat conclusions

1. Not “big” loss from flat tax
   - Optimal differentiated implies full policy maker information
   - Optimal differentiated does not account for administration/enforcement costs
2. Flat tax implies:
   - Lower taxes on low elasticity of demand goods (alcohol-tobacco, entertainment, rentals)
   - Higher taxes on high elasticity of demand goods (owned homes, services)

Exempted services frontiers comparison



Difference in exempted tax rates



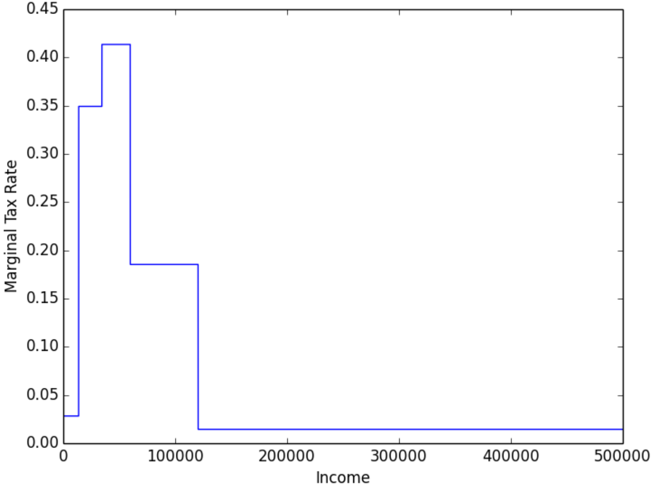
Exempted services conclusions

1. Minimal loss from exempting services
   - Less than 2% loss in total revenue
2. Zero services tax compensated by higher taxes on:
   - low elasticity goods at low revenues (alcohol-tobacco, entertainment, rentals)
   - high elasticity goods at high revenues (owned homes, taxed services)

Income Tax Example

- Households heterogeneous in:
  - ability (earning potential)
  - disutility of labor
- Income tax schedules τ piecewise linear marginal tax rates
  - six marginal tax rates
  - six income cutoffs
  - 12 dimensions of tax heterogeneity
- Now individual problem is nonconvex
- Requires huge database to solve

Simple optimal piecewise linear marginal tax rates



Summary

- Use big data techniques as solution method to theoretical models
- Ideal for:
  - High dimensional heterogeneity
  - Nonconvex optimization
  - Model and objective uncertainty

Future work

- Optimal differentiated Pigouvian gas tax with geographic pollution and congestion heterogeneity and households heterogeneity in commute times
- Optimal sales taxation in general equilibrium
- Optimal income taxation with dynamic household decisions