Impacts of Taxes on Firm Entry Rates along State Borders

Kevin D. Duncan, Georgeanne M. Artz, Peter F. Orazem [1]

Iowa State University

May 11th, 2016
Econometric Seminar

[1]This study was supported by a grant from the Charles Koch Foundation

Overview

- Contribution
- Variables of interest
- The estimation strategy
  - Sorting around the border
  - County matching procedure
  - Estimated equation
  - Additional covariates
  - Remaining problems
- The estimated effect
- Sensitivity analysis

Our contribution

- We use a larger array of top marginal tax rates than other papers in the literature
  - Property, income, sales, corporate, capital gains, workers compensation, unemployment insurance
- Papers traditionally use only a small subset of tax rates (e.g. income, corporate, sales), creating plausible omitted variable bias
- Some papers further don't use a non-endogenous proxy for effective tax rate, and we argue top marginal tax rate is the second best without firm characteristics
- The border discontinuity approach to this problem is not the novel contribution

Variable of Interest

- Want to estimate the impacts of a 1% increase in each of our different tax rates on firm start up rates
- Doesn't fit well into (L)ATE framework; firms sort over tax policy regimes, no unconfoundness
- Conditional logit and poisson models of sorting over all counties feature endogeneity of tax policy with respect to firm entry rates
- County-differencing strategy aims to control for policy endogeneity under a similar set of assumptions of sorting, while removing likelihood estimation ("Local sorting/entry model")
- End up estimat*ing* how a 1% increase in each of our different tax rates impacts *relative* firm start up rates

Sorting at the border

- Firms have face linear profit as a function of tax, regulatory, and location specific parameters.
- Firms enter into perfectly competitive markets, and so can only pick at best one period of arbitrage profits
- Conditional on picking to enter into a market around a state border, firms enter into the side of the border with the higher expected profits
- No matter what, the number of new firm start ups on either side of the border will never be almost random around with respect to the treatment (different tax rates)

Sorting at the border

- But do think of this as a sorting problem!
- We control for endogeneity of tax policy to firm entry rates by taking the difference in log firm start ups on states on either side of a state border; morphs dependent variable into relative firm start up rates, e.g. linear probability model
- As if we are pulling a random firm from those that have decided to enter into some market around the state border, and we estimate the relative probabilities of that firm choosing one side of the border over the other
- Much simpler criteria for identification, in particular, regular OLS identification requirements

Matching

- We use the Census Bureau's County Adjacency File (CAF) to match counties. The CAF orders states alphabetically, with counties alphabetically listed within. Adjacent counties are given similar treatment (e.g. first alphabetically by state, then alphabetically by county name).
- We assign the first column as "subject" counties, and the adjacent counties as "neighbors"
- Match each county in the subject column with every neighbor that is in another state.
- Repeat until we generated the first unique set of matches.

Matching: Example from the CAF

"Baldwin County, AL" 01003	"Monroe County, AL" 01099
"Baldwin County, AL" 01003	"Washington County, AL" 01129
"Baldwin County, AL" 01003	"Escambia County, FL" 12033
"Barbour County, AL" 01005	"Russell County, AL" 01113
"Barbour County, AL" 01005	"Clay County, GA" 13061
"Barbour County, AL" 01005	"Quitman County, GA" 13239

Matching: Visual Example

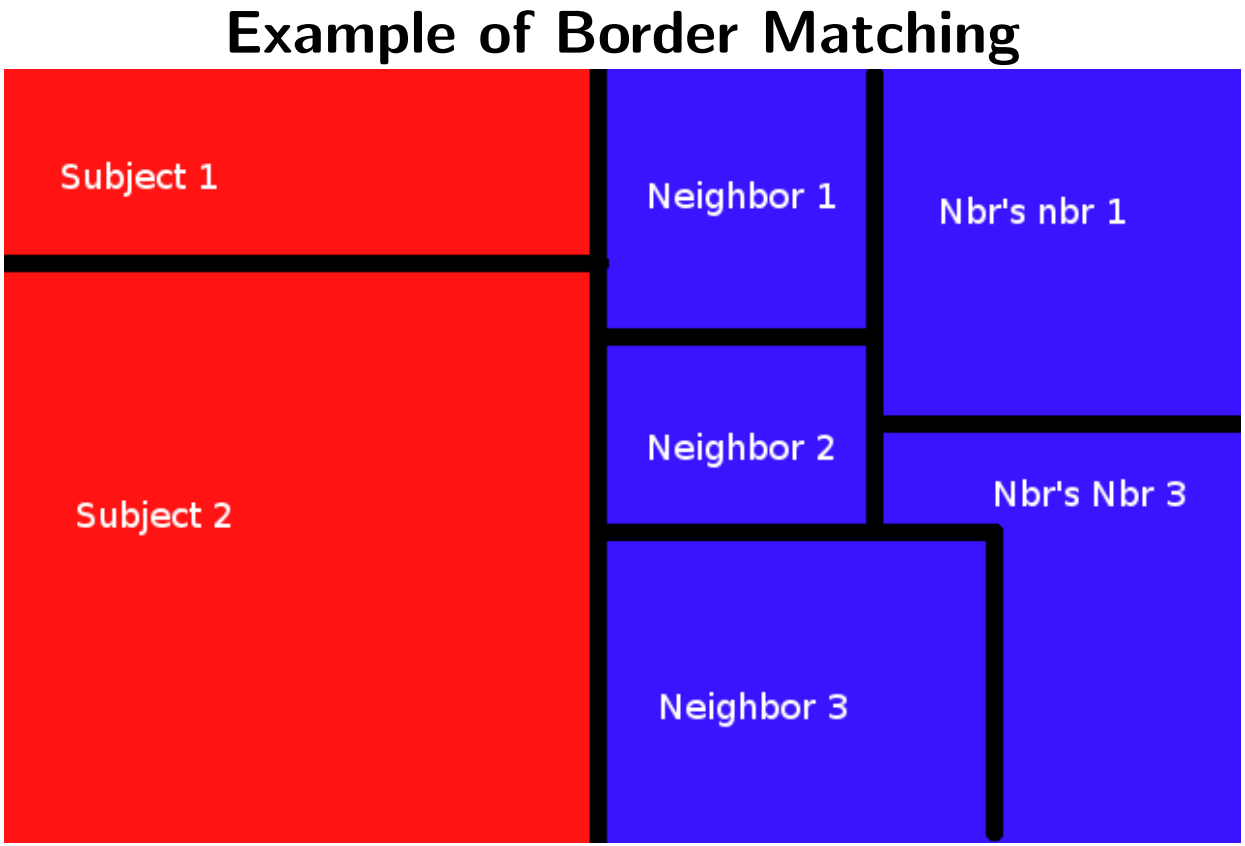

Figure : In this example Subject 1 would be only matched to Neighbor 1, while "Subject 2" would be paired with Neighbor 1-3. Similarly, when we broaden the bandwidth, Subject 1 would be matched with Nbr's Nbr 1, while Subject 2 would be paired with Nbr's Nbr 1 and 2.

Estimating Equation

- Imagine that we can approximate the log number of firms that enter into a county i in a state j in time t by the equation

$$\ln(n_{ijt}) = X_{it}\gamma + X_{jt}\beta + e_{ijt}$$

- Where X_{it} are location specific effects, X_{jt} are state specific effects, and e_{ijt} is a composite shock term
- Break up e_{ijt} into three components, where we are given a countable collection of markets $g = 1, ..., G$ that are sets of counties that may span different states.

$$e_{ijt} = \epsilon_t + \epsilon_{gt} + \epsilon_{ijt}$$

  where ϵ_t is shared across all counties in all states, ϵ_{gt} is shared across all counties in a particular market, and ϵ_{ijt} is a true idiosyncratic error.
- We want to estimate β.

Estimating Equation

- States may respond to both macroeconomic shocks along with federal governments, as well as market shocks that cover large areas of their territory. i.e. $\mathrm{E}(X'_{jt}e_t) \neq 0, \mathrm{E}(X'_{jt}e_{gt}) \neq 0$.
- Differencing between "small enough" areas on either side of the border controls for this endogeneity and allows us to estimate a differenced equation of the form,

$$\ln(n_{ijt}) - \ln(n_{i'j't}) = (X_{jt} - X_{j't})\beta + \epsilon_{ijt} - \epsilon_{i'j't}$$

- Estimate with POLS and clustered standard errors at the state-pair level

Additional Covariates

- log state spending per capita on infrastructure, education, and welfare
- optional state controls: real fuel price, percent unionized, percent manufacturing, percent high school education, population density
- optional scaled geographic amenities: percent water, January sunlight, January temperature, July temperature, July humidity, topology score
- State-pair specific fixed effects

Remaining Problems

- Properly capturing the relevant sorting variables along the border
  - No measure of county level agglomeration economies in the paper to date
  - Some counties are quite large, and no ability to create effective methods to control for this
- Omitted variable bias on not having county level tax and expenditure data
  - Recent articles show that counties should optimally try and counteract the state level tax differential
  - This shrinks results towards zero
- We cannot differentiate between pure new firms, firms that are part of chains that open up on both sides of the border, and firms that shut down on one side to open up on the other.

The Estimated Effect

Table : Regression Discontinuity Models for Total Firm Births

	births ratio	
Property Tax Difference	-0.371^{**}	-0.297^{**}
	(0.147)	(0.150)
Income Tax Difference	-0.085^{***}	-0.075^{***}
	(0.026)	(0.026)
Capital Gains Tax Difference	0.008	0.020
	(0.023)	(0.024)
Sales Tax Difference	-0.101^{***}	-0.087^{***}
	(0.030)	(0.032)
Corp Tax Difference	0.018	0.011
	(0.018)	(0.019)
Workers Comp Tax Difference	0.090	0.051
	(0.108)	(0.105)
Unemp. Tax Difference	0.012	-0.006
	(0.036)	(0.038)
log expend	yes	yes
state controls	yes	no
amenities	no	no
fixed effects	no	no
joint significance	yes	no
Observations	13,115	13,115
G	117	117
R^2	0.056	0.037

Sensitivity Analysis

- Not impose equality of the estimated effect on each side of the border
- States that share reciprocal agreements
- Sliding scale of urban v rural areas
- Fixed effects for each state
- Cross-sectional analysis for each year
- Subsample estimates for 2-digit NAICS sectors
- Extend the distance between each matched pair

Thank you!