

Departamento de Economia  
 Universidade de São Paulo  
 EAE-6029 - Final Exam  
 Prof.: Pedro Forquesato  
 May 11, 2023

Name: \_\_\_\_\_ Nº USP: \_\_\_\_\_

**Instructions:** The exam is **individual and closed book**. The exam is worth 10 points. The allotted time is **3 hours**. All answers must be complete and well explained to be worth points. Answers that are tangential to the question will be ignored. The exam must be completed using a pen.

1. (1 ponto) Consider this estimating equation from the paper Harding et al (2012) (mentioned in class):

$$P_{uijt} = \beta_0 + \beta_1 \tau_{jt}^h + \beta_2 (\tau^h - \tau^b)_{jt} + \beta_3 \ln D_{ijt} + \beta_4 (\tau^h - \tau^b)_{jt} \cdot \ln D_{ijt} + \theta X_i + \delta_j + \psi_t + \alpha_u + \epsilon_{uijt}$$

Where  $P$  is the price of product  $u$  bought by individual  $i$  at store  $j$  and time  $t$ ,  $\tau^h$  is the home state tax,  $\tau^b$  is the neighbor state tax,  $D$  is the border distance, and  $X$  is a vector of controls. The results are presented in the table below.

TABLE 3—OLS ESTIMATES OF THE EFFECT OF EXCISE TAXES ON CIGARETTE PRICES BY  
 DISTANCE TO LOWER-TAX BORDERS

Independent variable	(1)	(2)	(3)
<i>Panel A. Distance from home census tract</i>			
Excise tax (cents)	0.798** (0.077)	0.807** (0.073)	0.856** (0.043)
Tax difference with nearest lower-tax state	-0.514** (0.180)	-0.587** (0.176)	-0.362** (0.091)
Log distance to nearest lower-tax state	-3.412 (2.406)	-4.490* (2.352)	-1.489 (1.005)
Log distance-tax difference interaction for nearest lower-tax state	0.139** (0.042)	0.153** (0.041)	0.076** (0.020)
Demographic controls:	No	Yes	Yes
UPC fixed effects:	No	No	Yes

What is the effect of a rise in cigarette taxes at the home state  $\tau^h$  (and therefore the home to border state differential  $\tau^h - \tau^b$ ) on the price of the product? Be *entirely* precise.

2. (2 pontos) In a paper presented in our department seminar April 20th, 2023, Sophie Mattes, professor of FGV/EPGE, and co-authors study the effect of better local longevity on local house prices, controlling for many local amenities, like violent crime, air pollution among others. (Other details of their empirical strategy are not relevant here.)
- (a) What would happen if they did not add these controls? Be precise, with exact examples. If there is a bias, be clear about its direction (and why).

- (b) In the seminar, Sophie claimed “We don’t control for hospital quality, because that would be a bad control”. Is hospital quality a bad control? What about a collider? (As mentioned in class, the terminology here is ambiguous.) Explain precisely using a directed acyclic graph and giving real world examples.
3. (2 pontos) In a paper that follows from his undergraduate thesis, Pedro Rodrigues (and me, 2023) argue that the staggered introduction of biometric registration in Brazilian elections *caused* a reduction in the number of registered voters, with no effect on the number of votes cast on ballot. We estimate this using a difference-in-differences design, comparing municipalities that adopted the biometric with municipalities that did not adopt it during our sample period.
- (a) What are the identification assumptions of this empirical exercise? Give a (precise, real world) example of when it would be violated.
- (b) In the paper, our observation are at the level of municipality (4500) and election (6), but we analyze elections for state governors (24 states). One problem we find is that when we cluster at the municipality level (and election year), pre-trends are statistically significant, while clustering on state level solves this issue (while our main results remain significant). Discuss (econometrically) the pros and cons of each approach. Which one you find more reasonable?
4. (1 ponto) Consider  $\tilde{Y} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ . Find the OLS coefficient of a regression of  $\tilde{Y}$  on  $\mathbf{X}$ .
5. (1 ponto) In the standard IV structural model  $Y = X'\beta + e$  and  $X = Z'\Gamma + u$ , consider the usual control function  $e = u'\eta + v$ . Show (algebraically) that  $\mathbb{E}[Xv] = 0$ .
6. (1 ponto) Is  $\mathbb{E}[\epsilon_{it} | X_{it}] = 0$  enough for the fixed effects estimator to be consistent for  $\beta$ ? Show it.
7. (1 ponto) Consider a linear regression model  $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + e$ , with estimates  $\hat{\alpha} = 0$ ,  $\hat{\beta}_1 = 3$ , and  $\hat{\beta}_2 = 10$ , and we want to test the (three-dimensional) null hypothesis:  $\alpha = 0$ ,  $\beta_1^2 = \beta_2$  and  $\beta_2 - \alpha = 10$ .

The variance-covariance matrix is:

$$\Sigma = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Find the Wald test statistic and its (precise) asymptotic distribution.

8. (1 ponto) We saw in class that OLS is the sample analog of the linear projection, which is the best linear predictor. Is the 2SLS also the best linear predictor? What maximization problem does it solve? Show both.