Homework 3: Submission 2

Research Methods, Spring 2024

Leila Mulveny

Homework 3: Repository

Summarize The Data

1. Present a bar graph showing the proportion of states with a change in their cigarette tax in each year from 1970 to 1985.

See Figure 1:

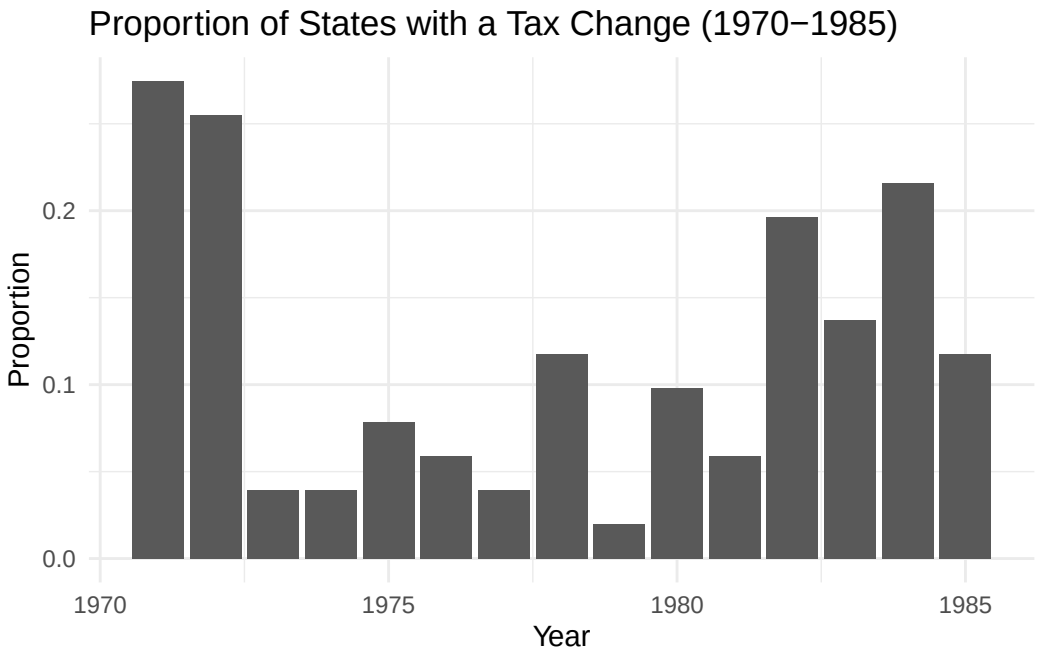

Figure 1: Proportion of States With a Tax Change in Their Cigarette Tax from 1970-1985

2. Plot on a single graph the average tax (in 2012 dollars) on cigarettes and the average price of a pack of cigarettes from 1970 to 2018.

See Figure 2:

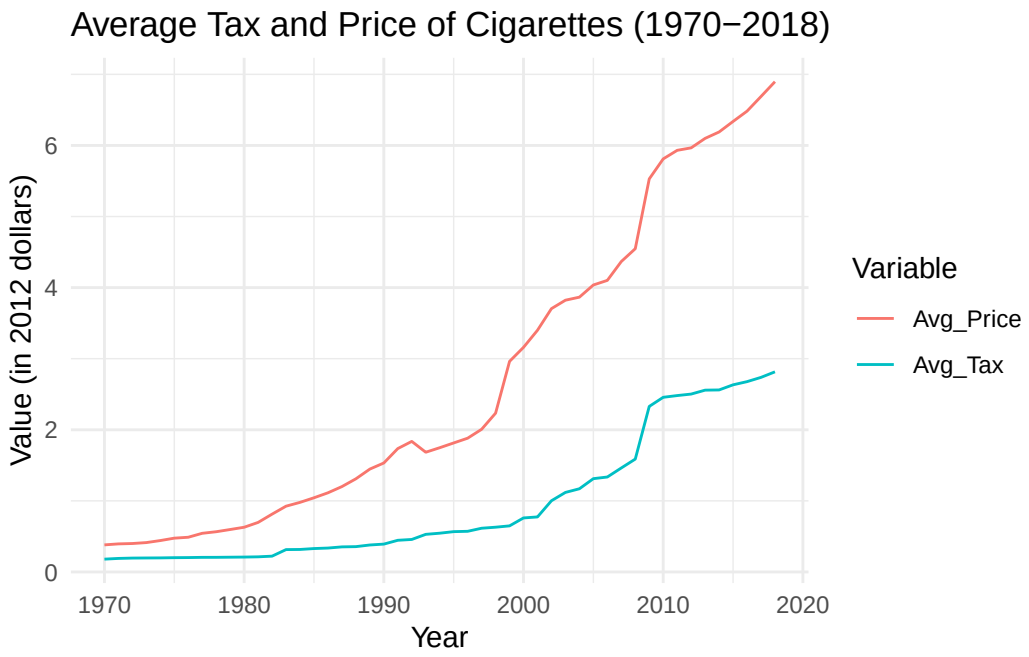

Figure 2: Average Tax (in 2012 dollars) on Cigarettes and the Average Price of a Pack of Cigarettes from 1970-2018

3.Identify the 5 states with the highest increases in cigarette prices (in dollars) over the time period. Plot the average number of packs sold per capita for those states from 1970 to 2018.

See Figure 3:

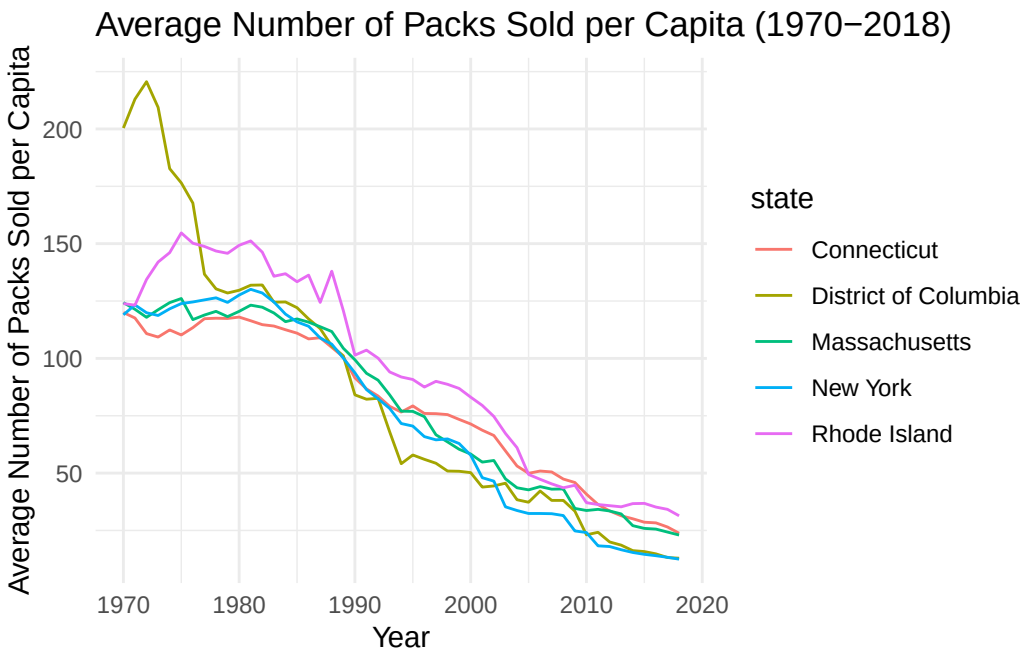

Figure 3: Average Number of Packs Sold per Capita for the Top Five States with the Highest increases in Cigarette Prices (in dollars) from 1970-2018

4. Identify the 5 states with the lowest increases in cigarette prices (in dollars) over the time period. Plot the average number of packs sold per capita for those states from 1970 to 2018.

See Figure 4:

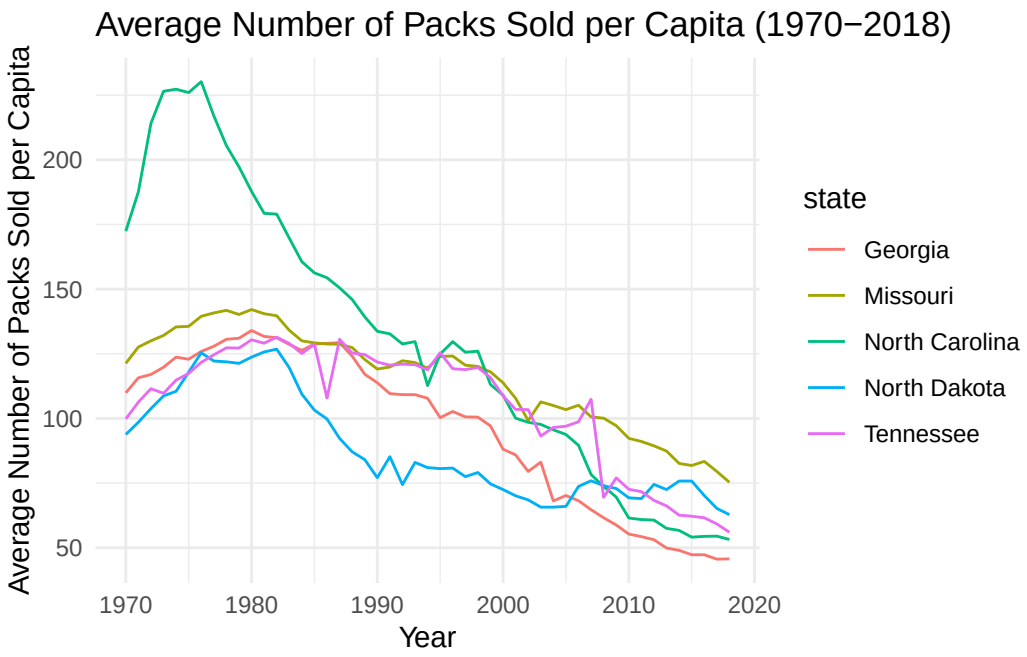

Figure 4: Average Number of Packs Sold per Capita for the Bottom Five States with the Lowest increases in Cigarette Prices (in dollars) from 1970-2018

5.Compare the trends in sales from the 5 states with the highest price increases to those with the lowest price increases.

See Figure 5:

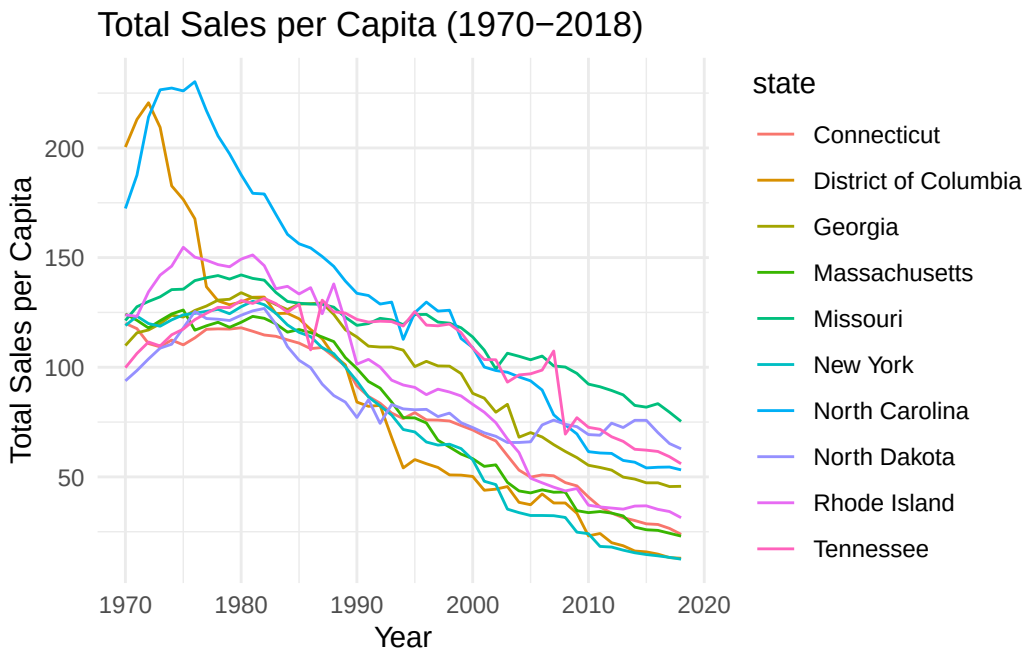

Figure 5: Comparison of the Trends in Sales From the 5 States with the Highest Price Increase to Those with the Lowest Price Increases

Both the states with the highest price increase and those with the lowest price increase demonstrate a similar trend, which is a decrease from 1970 - 2020. Of the high states, DC sales peak a little after 1970 while the other 4 states peak closer to 1980 before starting to deacrease. Of the low states, North Carolina peaks in 1975 and then steeply falls off, while the other 4 states softly peak in 1980 before decreasing. Generally the sales trend is the same (decrease after 1980) in even the highest and lowest price change states.

Estimate ATEs

Now let's work on estimating a demand curve for cigarettes. Specifically, we're going to estimate the price elasticity of demand for cigarettes. When explaining your findings, try to limit your discussion just to a couple of sentences..

6. Focusing only on the time period from 1970 to 1990, regress log sales on log prices to estimate the price elasticity of demand over that period. Interpret your results.

See @tab-pe

Table 1: Regression Output of Log(Sales) on Log(Prices)

Label	Coefficient	Standard.Error
Beta 0	5.4273812	0.0297515
Beta 1	-0.8094384	0.0383656

The results displayed in @tab-pe show a coefficient of -0.809 which is the estimated price elasticity of demand. It means that a 1% increase in cigarette price is associated with a decrease of about 0.809% in sales. This is result makes sense: as a normal good quantity demanded of cigarettes should decrease as price increases.

7. Again limiting to 1970 to 1990, regress log sales on log prices using the total (federal and state) cigarette tax (in dollars) as an instrument for log prices. Interpret your results and compare your estimates to those without an instrument. Are they different? If so, why?

See (@tab-1-iv):

Table 2: Regression Output of Log(Sales) on Log(Prices) with Cigarette Tax as an Instrumental Variable

Label	Coefficient	Standard.Error
Beta 0	5.4167966	0.0544940
Beta 1	-0.7955235	0.0712351

The results displayed in @tab-1-iv show a coefficient of about -0.8 which is the estimated price elasticity of demand. It means that a 1% increase in cigarette price is associated with a decrease of about 0.8% in sales. Again, this is intuitive because an increase in price should result in a decrease in quantity demand. This beta is slightly smaller in absolute value than the beta estimated in @tab-1-iv, we can assume that the coefficient for @tab-1-pe is slightly larger due to unaccounted for endogeneity between price and the error term, which the instrument accounts for in @tab-1-iv.

8. Show the first stage and reduced-form results from the instrument.

See (@tab-stage-1):

Table 3: First Stage Regression Output of Log(Sales) on Log(Prices) with Cigarette Tax as an Instrumental Variable

Label	Coefficient	Standard.Error
Beta 0	0.8396457	0.0054213
Beta 1	0.2600601	0.0124425

See (@tab-reduced-1):

Table 4: Reduced Form Regression Output of Log(Sales) on Log(Prices) with Cigarette Tax as an Instrumental Variable

Label	Coefficient	Standard.Error
Beta 0	4.7488387	0.0092015
Beta 1	-0.2068839	0.0211186

Estimate ATEs

Question 9 9. Repeat questions 6-8 focusing on the period from 1991 to 2015.

9.1 Focusing only on the time period from 1991 to 2015, regress log sales on log prices to estimate the price elasticity of demand over that period. Interpret your results.

See (@tab-2-pe):

Table 5: Regression Output of Log(Sales) on Log(Prices) from 1991-2015

Label	Coefficient	Standard.Error
Beta 0	5.6599553	0.0363844
Beta 1	-0.9968136	0.0246921

The results in @tab-2-pe display a coefficient for "ln_price" of -0.99. It means that a 1% increase in price is associated with a decrease of about 0.99% in sales, which is nearly 1-1 (price elasticity of a unit elastic good).

9.2 Again limiting to 1991 to 2015, regress log sales on log prices using the total (federal and state) cigarette tax (in dollars) as an instrument for log prices. Interpret your results and compare your estimates to those without an instrument. Are they different? If so, why?

See (@tab-2-iv):

Table 6: Regression Output of Log(Sales) on Log(Prices) with Cigarette Tax as an Instrumental Variable from 1991-2015

	model_2
Dependent Var.:	**ln_sales**
Constant	5.880*** (0.0408)
ln_price	-1.150*** (0.0278)
S.E. type	IID
Observations	1,275
R2	0.54817
Adj. R2	0.54782

The results displayed in @tab-2-iv show a coefficient of -1.15 while @tab-2-pe yeilds a coefficient of -0.99. Both are negative, but the coefficient of -1.15 is much larger in absolute value than -0.99. This suggests that when accounting for the potential endogeneity of price (using "ln_total_tax" as an instrument), a 1% increase in price is associated with a larger decrease in sales (about 1.15% instead of 0.99%). The differences could be due to the endogeneity of "ln_price", which @tab-2-iv tries to correct for by using "ln_total_tax" as an instrument.

9.3 Show the first stage and reduced-form results from the instrument.

See (@tab-stage-2):

Table 7: First Stage Regression Output of Log(Sales) on Log(Prices) with Cigarette Tax as an Instrumental Variable from 1991-2015

Label	Coefficient	Standard.Error
Beta 0	1.3150730	0.0043859
Beta 1	0.5135503	0.0069225

See (@tab-reduced-2):

Table 8: Reduced Form Regression Output of Log(Sales) on Log(Prices) with Cigarette Tax as an Instrumental Variable from 1991-2015

Label	Coefficient	Standard.Error
Beta 0	4.3674179	0.0084395
Beta 1	-0.5906258	0.0133205

10. Compare your elasticity estimates from 1970-1990 versus those from 1991-2015. Are they different? If so, why?.

The estimated price elasticity of demand is -0.8 in @tab-1-iv and -1.15 @tab-2-iv. While both estimates are negative, @tab-2-iv estimates a larger price elasticity of demand in absolute value. Cigarette smoking was much more prevelant from 1970-1990, and because of the addictive quality of cigarettes we would expect demand to more more inelastic in 1970-1990 than compared to 1991-2015. This explains why the price elasticity of demand is smaller in absolute value from 1970-1990 (its behaving more so as a necessary good).