Submission1-HW3

Research Methods, Spring 2024

Caroline Hansen

https://github.com/carolinezhansen/tobacco/tree/main

Answers for Homework 3: Submission 1

Problem 1

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

```
      [,1]
 [1,]  0.7
 [2,]  1.9
 [3,]  3.1
 [4,]  4.3
 [5,]  5.5
 [6,]  6.7
 [7,]  7.9
 [8,]  9.1
 [9,] 10.3
[10,] 11.5
[11,] 12.7
[12,] 13.9
[13,] 15.1
[14,] 16.3
[15,] 17.5
[16,] 18.7
```

Problem 2

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.

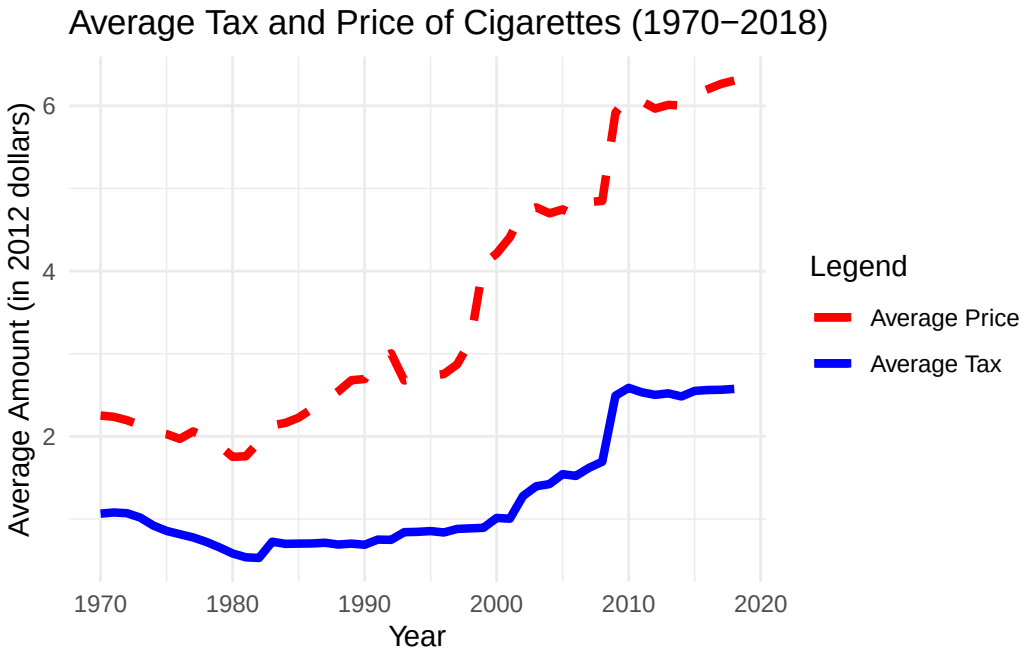

Figure 1: Average Tax and Average Price of Cigarettes from 1970 to 2018

Problem 3

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.

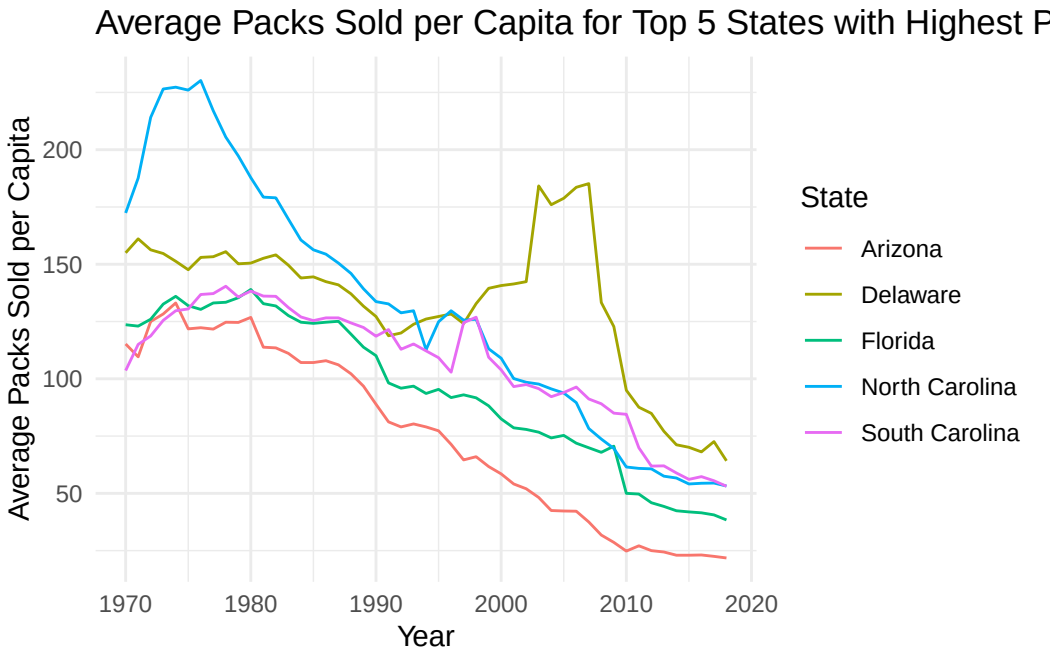

Figure 2: Average Packs Sold per Capita for Top 5 States with Highest Price Increases

Problem 4

4. 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.

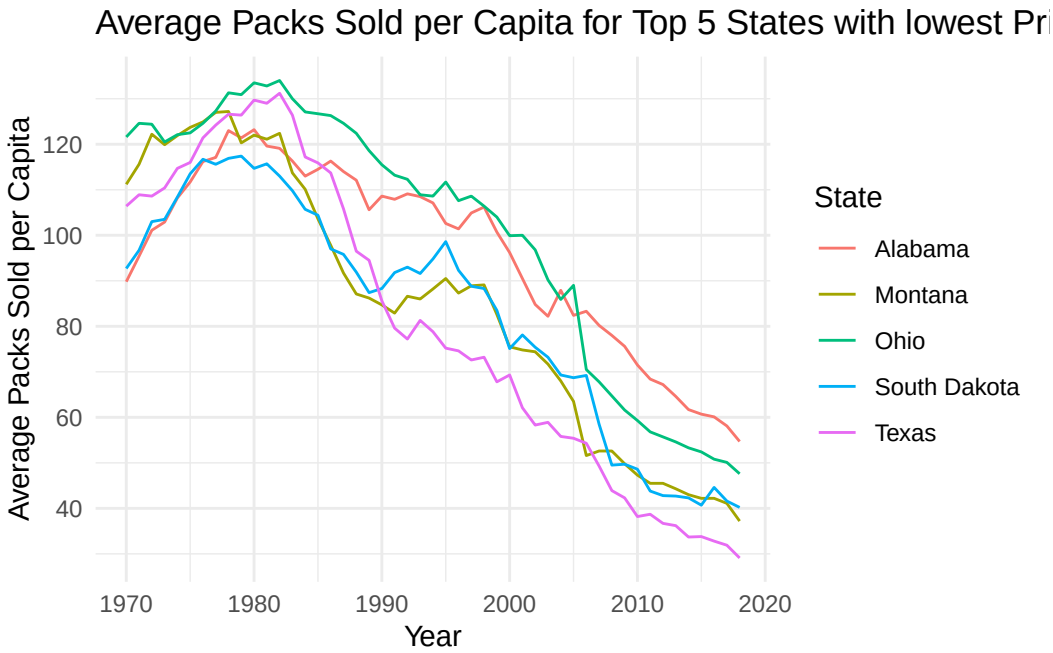

Figure 3: Average Packs Sold per Capita for Bottom 5 States with Lowest Price Increases

Problem 5

5. 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.

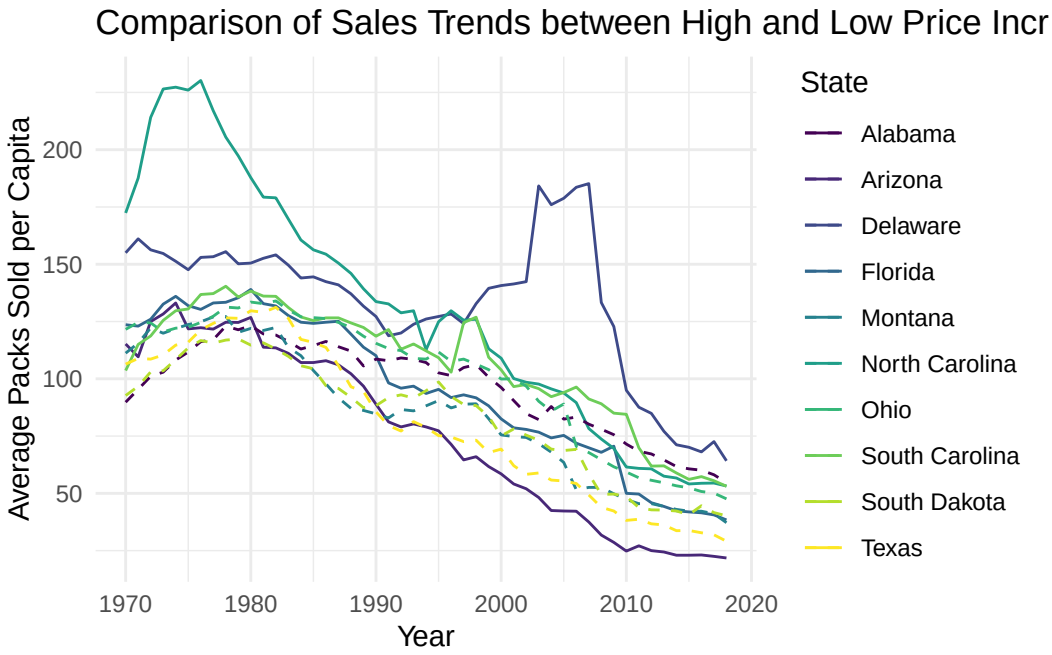

Figure 4: Comparison of Sales Trends between High and Low Price Increase States

Problem 6

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.

```
OLS estimation, Dep. Var.: ln_sales
Observations: 1,071
Standard-errors: IID
             Estimate Std. Error  t value  Pr(>|t|)
(Intercept)  4.750402   0.008116 585.3207 < 2.2e-16 ***
ln_price    -0.171540   0.013829 -12.4039 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.210546   Adj. R2: 0.125
```

Problem 7

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?

```
TSLS estimation, Dep. Var.: ln_sales, Endo.: ln_price, Instr.: ln_total_tax
Second stage: Dep. Var.: ln_sales
Observations: 1,224
Standard-errors: IID
             Estimate Std. Error  t value  Pr(>|t|)
(Intercept)   5.208117   0.026904 193.5834 < 2.2e-16 ***
fit_ln_price -0.804393   0.021315 -37.7388 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.311225   Adj. R2: 0.488176
F-test (1st stage), ln_price: stat = 3,702.1, p < 2.2e-16, on 1 and 1,222 DoF.
                  Wu-Hausman: stat =   274.0, p < 2.2e-16, on 1 and 1,221 DoF.
```

Problem 8

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

```
OLS estimation, Dep. Var.: ln_price
Observations: 1,071
Standard-errors: IID
              Estimate Std. Error   t value  Pr(>|t|)
(Intercept)  -0.503589   0.020736 -24.28593 < 2.2e-16 ***
ln_total_tax -0.411813   0.043812  -9.39948 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.447517   Adj. R2: 0.075474

OLS estimation, Dep. Var.: ln_price
Observations: 1,224
Standard-errors: IID
             Estimate Std. Error  t value  Pr(>|t|)
(Intercept)  1.079743   0.007105 151.9773 < 2.2e-16 ***
ln_total_tax 0.718191   0.011804  60.8451 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.239976   Adj. R2: 0.751631
```

Problem 9

9.# 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?

```
OLS estimation, Dep. Var.: ln_sales
Observations: 1,224
Standard-errors: IID
             Estimate Std. Error  t value  Pr(>|t|)
(Intercept)  5.024310   0.023058 217.8955 < 2.2e-16 ***
ln_price    -0.650076   0.017947 -36.2225 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.302217   Adj. R2: 0.517377


TSLS estimation, Dep. Var.: ln_sales, Endo.: ln_price, Instr.: ln_total_tax
Second stage: Dep. Var.: ln_sales
Observations: 1,224
Standard-errors: IID
              Estimate Std. Error  t value  Pr(>|t|)
(Intercept)   5.208117   0.026904 193.5834 < 2.2e-16 ***
fit_ln_price -0.804393   0.021315 -37.7388 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.311225   Adj. R2: 0.488176
F-test (1st stage), ln_price: stat = 3,702.1, p < 2.2e-16, on 1 and 1,222 DoF.
                  Wu-Hausman: stat =   274.0, p < 2.2e-16, on 1 and 1,221 DoF.


OLS estimation, Dep. Var.: ln_price
Observations: 1,224
Standard-errors: IID
             Estimate Std. Error  t value  Pr(>|t|)
(Intercept)  1.079743   0.007105 151.9773 < 2.2e-16 ***
ln_total_tax 0.718191   0.011804  60.8451 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.239976   Adj. R2: 0.751631


OLS estimation, Dep. Var.: ln_price
Observations: 1,224
```

```
Standard-errors: IID
             Estimate Std. Error  t value  Pr(>|t|)
(Intercept)  1.079743   0.007105 151.9773 < 2.2e-16 ***
ln_total_tax 0.718191   0.011804  60.8451 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.239976   Adj. R2: 0.751631
```