Homework 3

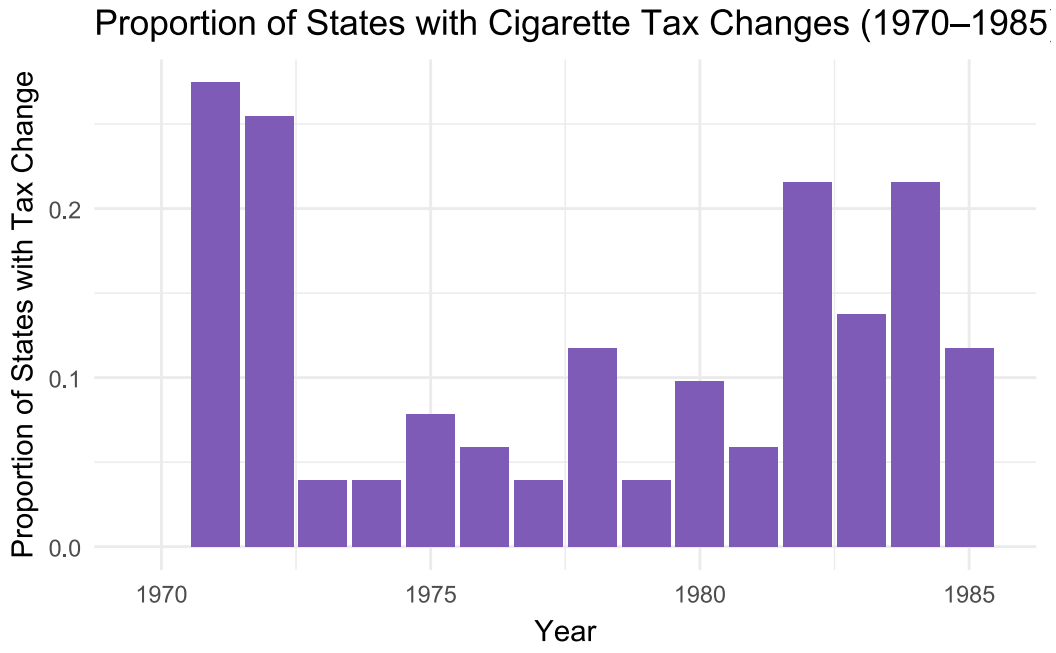
Nikhita Gandhe

2025-03-18

Please find the link to my GitHub repository here: https://github.com/nsgand2/ECON-470-HW3.git

Question 1

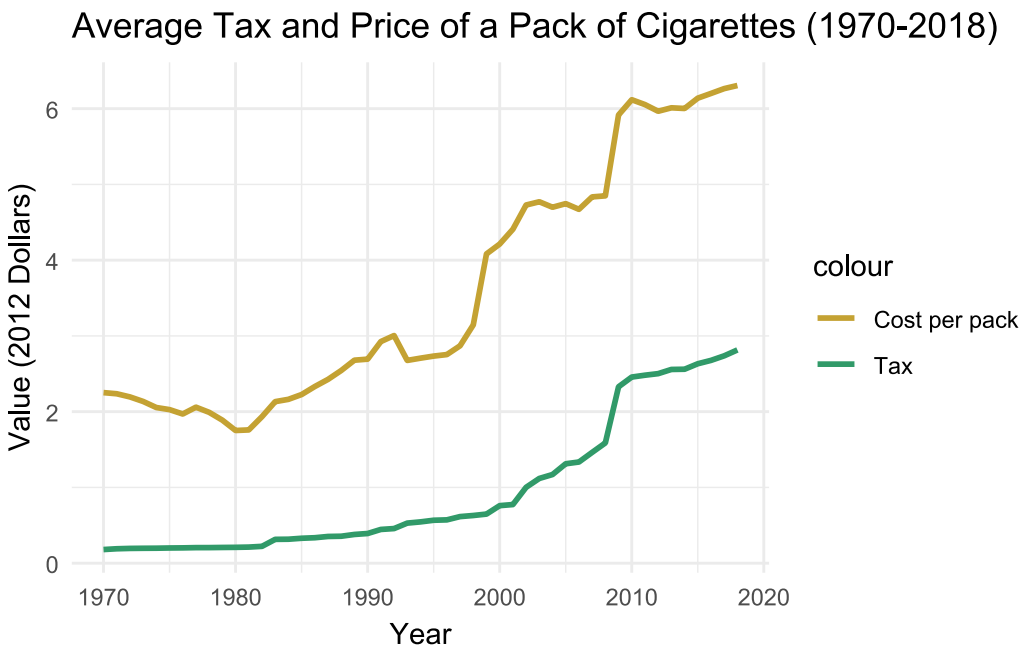
Present a bar graph showing the proportion of states with a change in their cigarette tax in each year from 1970 to 1985.



Question 2

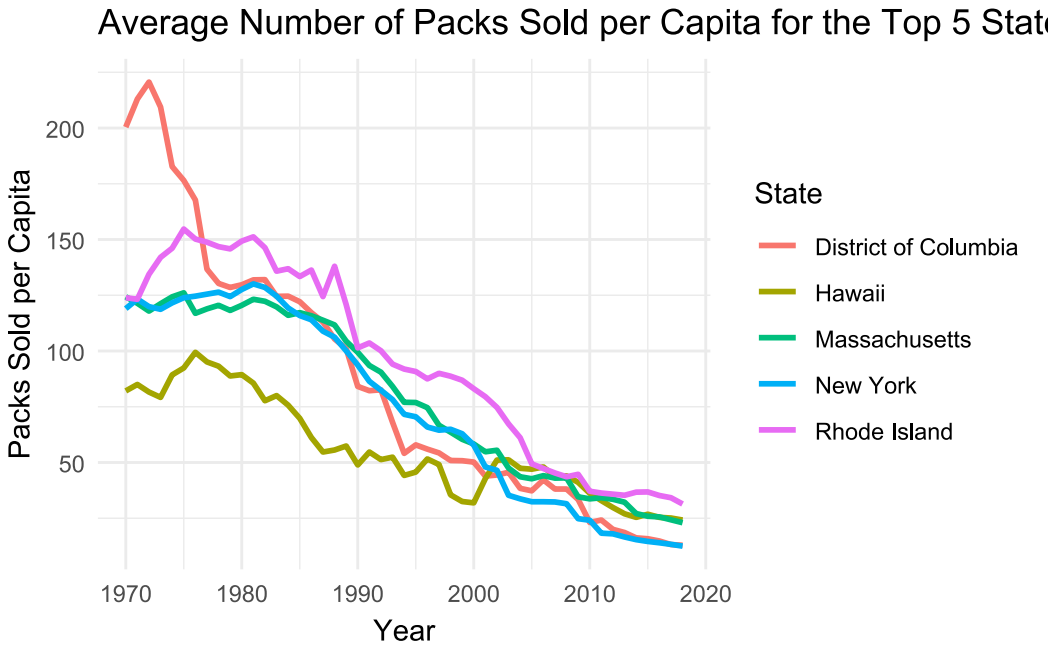
Plot on a single graph the average tax (in 2012 dollars) on cigarettes and the average price of a pack of cigarettes from 1970 to 2018.

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
ℹ Please use `linewidth` instead.
```



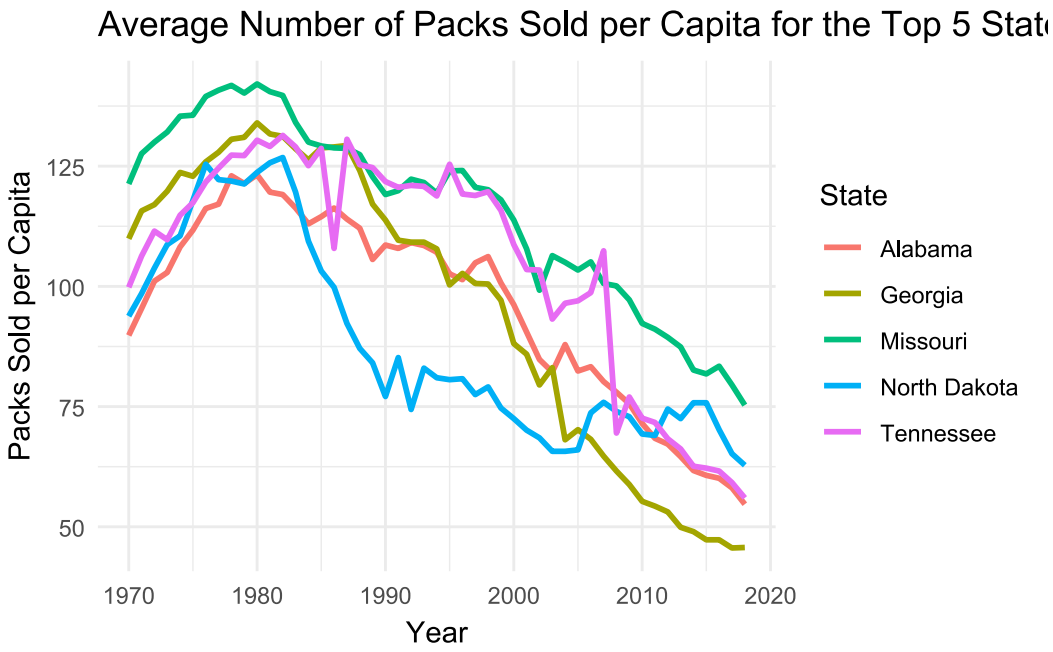
Question 3

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.



Question 4

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



Question 5

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

The number of cigarettes sold has significantly declined in both the states with the highest and lowest price increases, reflecting broader cultural shifts away from smoking. Missouri and North Dakota contribute to a slight rise in the average cost per pack among the lower-priced states. Currently, both groups see approximately 50 packs sold per capita. However, the lower-priced states exhibit greater variability over time, with North Carolina reaching around 250 packs sold per capita in the mid-1970s.

Question 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.

```
Call:
lm(formula = ln_sales ~ ln_price_cpi, data = cig.data_1970_1990)

Residuals:
```

```
     Min       1Q   Median       3Q      Max
-0.68335 -0.08598 -0.00284  0.08778  0.83516

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.42738    0.02975   182.4   <2e-16 ***
ln_price_cpi -0.80944    0.03837   -21.1   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1894 on 1069 degrees of freedom
Multiple R-squared:  0.294, Adjusted R-squared:  0.2933
F-statistic: 445.1 on 1 and 1069 DF,  p-value: < 2.2e-16
```

Question 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?

```

Call:
lm(formula = ln_sales ~ pricehat, data = cig.data_1970_1990)

Residuals:
     Min       1Q   Median       3Q      Max
-0.86239 -0.09798  0.00549  0.09359  0.95094

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.41679    0.06212  87.196   <2e-16 ***
pricehat    -0.79552    0.08121  -9.796   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2159 on 1069 degrees of freedom
Multiple R-squared:  0.08238,   Adjusted R-squared:  0.08152
F-statistic: 95.97 on 1 and 1069 DF,  p-value: < 2.2e-16
```

With a coefficient of 0.5, cigarette demand appears inelastic, as a 1% increase in price results in only a 0.5% change in quantity demanded. This estimate differs significantly from the OLS elasticity results, likely due to endogeneity issues in the OLS model, which the instrumental variables approach helps address.

Question 8

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

```
Call:
lm(formula = ln_price_cpi ~ ln_total_tax, data = (cig.data_1970_1990 %>%
    filter(Year >= 1970 & Year <= 1990)))

Residuals:
     Min       1Q   Median       3Q      Max
-0.23046 -0.09207 -0.02919  0.08019  0.48675

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.839182   0.005406   155.2   <2e-16 ***
ln_total_tax 0.260060   0.012443    20.9   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1272 on 1069 degrees of freedom
Multiple R-squared:  0.2901,    Adjusted R-squared:  0.2894
F-statistic: 436.8 on 1 and 1069 DF,  p-value: < 2.2e-16
```

```
Call:
lm(formula = ln_sales ~ pricehat, data = cig.data_1970_1990)

Residuals:
     Min       1Q   Median       3Q      Max
-0.86239 -0.09798  0.00549  0.09359  0.95094

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.41679    0.06212  87.196   <2e-16 ***
pricehat    -0.79552    0.08121  -9.796   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2159 on 1069 degrees of freedom
Multiple R-squared:  0.08238,   Adjusted R-squared:  0.08152
F-statistic: 95.97 on 1 and 1069 DF,  p-value: < 2.2e-16
```

Question 9

Repeat questions 1-3 focusing on the period from 1991 to 2015.

OLS estimates

```
Call:
```

```
lm(formula = ln_sales ~ ln_price_cpi, data = (cig.data_1991_2015 %>%
    filter(Year >= 1991 & Year <= 2015)))

Residuals:
     Min       1Q   Median       3Q      Max
-0.92230 -0.17004  0.00664  0.17869  1.10282

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.65995    0.03638  155.56   <2e-16 ***
ln_price_cpi -0.99681    0.02469  -40.37   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.296 on 1273 degrees of freedom
Multiple R-squared:  0.5614,    Adjusted R-squared:  0.5611
F-statistic:  1630 on 1 and 1273 DF,  p-value: < 2.2e-16
```

As price increases by 1%, quantity demanded decreases by 66%.

IV estimates

```

Call:
lm(formula = ln_sales ~ pricehat, data = (cig.data_1991_2015 %>%
    filter(Year >= 1991 & Year <= 2015)))

Residuals:
     Min       1Q   Median       3Q      Max
-0.90878 -0.15465  0.01119  0.15334  1.16925

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.87986    0.03803  154.60   <2e-16 ***
pricehat    -1.15008    0.02594  -44.34   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2802 on 1273 degrees of freedom
Multiple R-squared:  0.607, Adjusted R-squared:  0.6067
F-statistic:  1966 on 1 and 1273 DF,  p-value: < 2.2e-16
```

With a coefficient of −0.8, cigarette demand remains inelastic, as a 1% increase in price leads to only a 0.8% decrease in quantity demanded. While this estimate differs from the OLS elasticity results, the contrast is less pronounced compared to previous discrepancies between OLS and instrumental variable estimates.

2SLS estimates

```
Call:
lm(formula = ln_price_cpi ~ ln_total_tax, data = (cig.data_1991_2015 %>%
    filter(Year >= 1991 & Year <= 2015)))

Residuals:
     Min       1Q   Median       3Q      Max
-0.36750 -0.09020  0.00725  0.08241  0.45045

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.314161   0.004390  299.33   <2e-16 ***
ln_total_tax 0.513550   0.006922   74.19   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1456 on 1273 degrees of freedom
Multiple R-squared:  0.8121,    Adjusted R-squared:  0.812
F-statistic:  5504 on 1 and 1273 DF,  p-value: < 2.2e-16
```

```
Call:
lm(formula = ln_sales ~ pricehat, data = (cig.data_1991_2015 %>%
    filter(Year >= 1991 & Year <= 2015)))

Residuals:
     Min       1Q   Median       3Q      Max
-0.90878 -0.15465  0.01119  0.15334  1.16925

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.87986    0.03803  154.60   <2e-16 ***
pricehat    -1.15008    0.02594  -44.34   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2802 on 1273 degrees of freedom
Multiple R-squared:  0.607, Adjusted R-squared:  0.6067
F-statistic:  1966 on 1 and 1273 DF,  p-value: < 2.2e-16
```

Question 10

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

1970-1990 The results indicate a positive elasticity, meaning that an increase in price corresponds with higher sales—an uncommon pattern in most markets. This anomaly could stem from external

factors such as tax hikes, policy changes (e.g., tobacco regulations), or broader structural shifts. One possible explanation is that higher taxes may have made cigarettes appear more exclusive or prestigious, leading to increased consumption despite rising prices.

1991-2015 In contrast, this period exhibits the expected negative elasticity, where higher cigarette prices correspond with lower sales. This aligns with standard economic theory and consumer behavior, where increased costs typically discourage consumption. The shift likely reflects stronger public health campaigns, smoking bans, and greater awareness of the risks associated with smoking, reinforcing the expected inverse relationship between price and demand.