Bhasin-S-hwk3-1

Sachi Bhasin
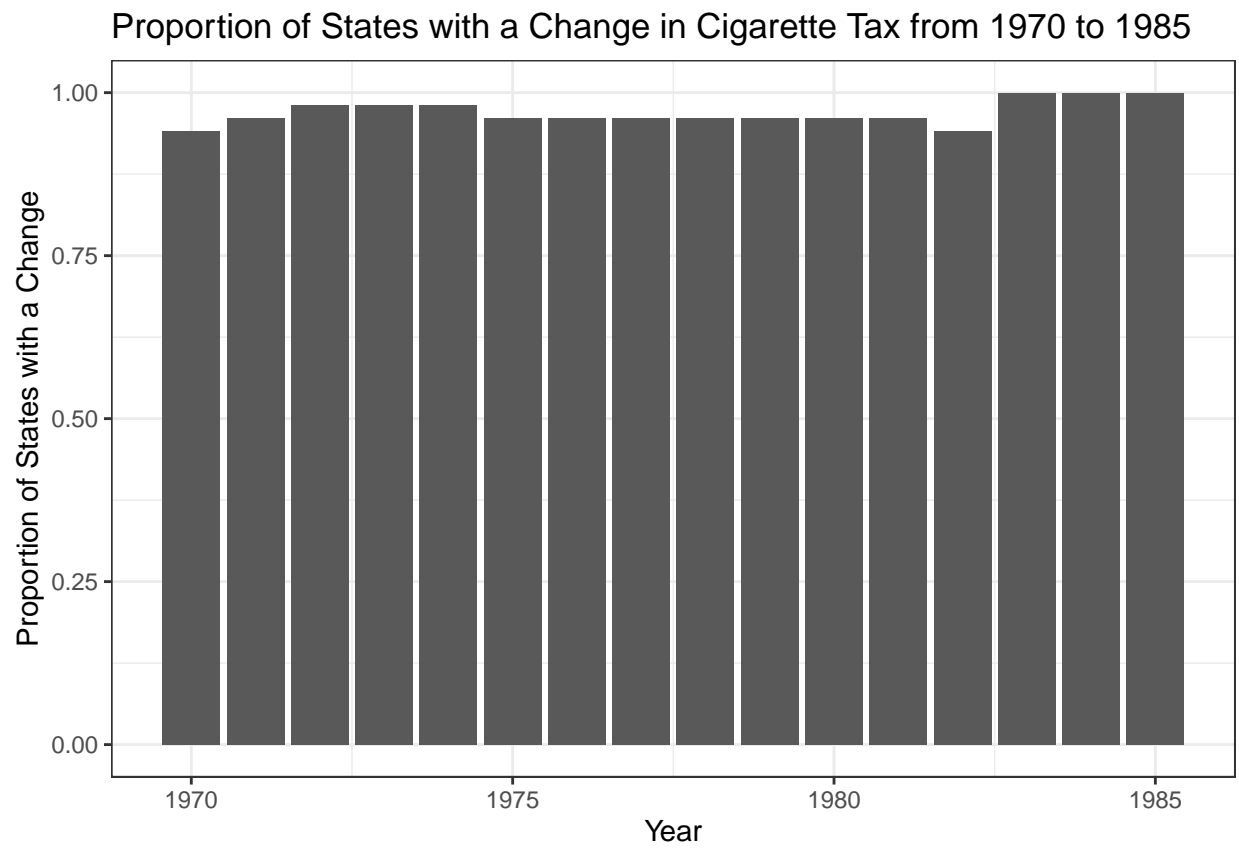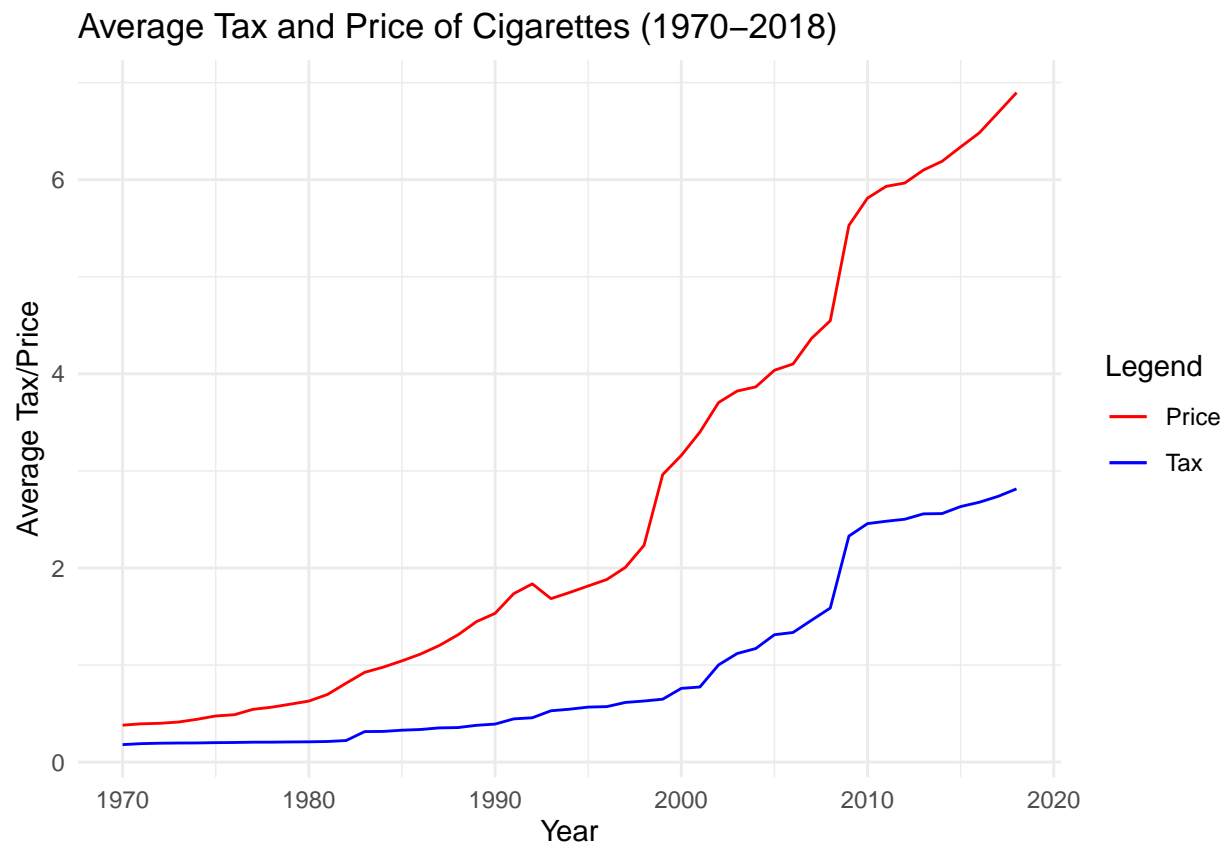
2023-03-12

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

Proportion of States with a Change in Cigarette Tax from 1970 to 1985

Average Tax and Price of Cigarettes (1970–2018)

Question 3

```
## # A tibble: 5 x 2
##   state                price_change
##   <chr>                       <dbl>
## 1 New York                     9.93
## 2 District of Columbia         9.54
## 3 Connecticut                  9.30
## 4 Rhode Island                 9.20
## 5 Massachusetts                9.18
```
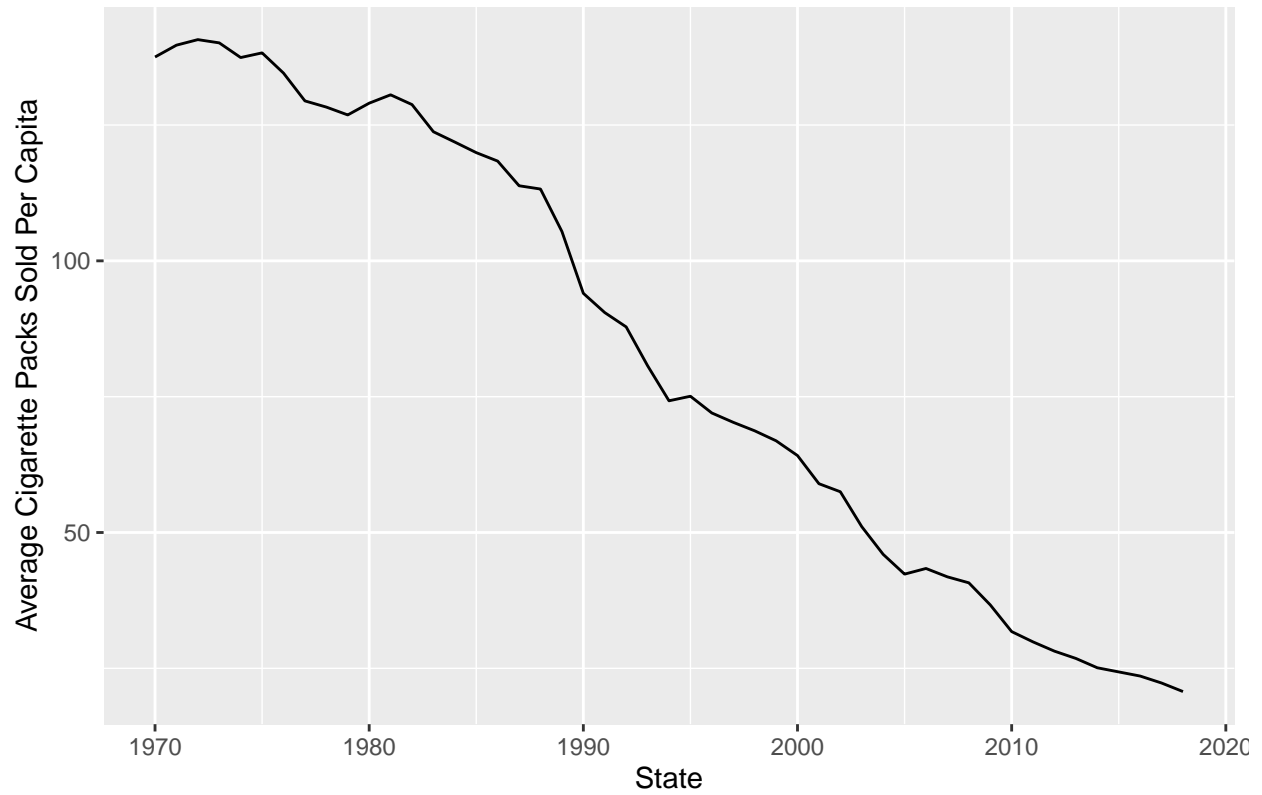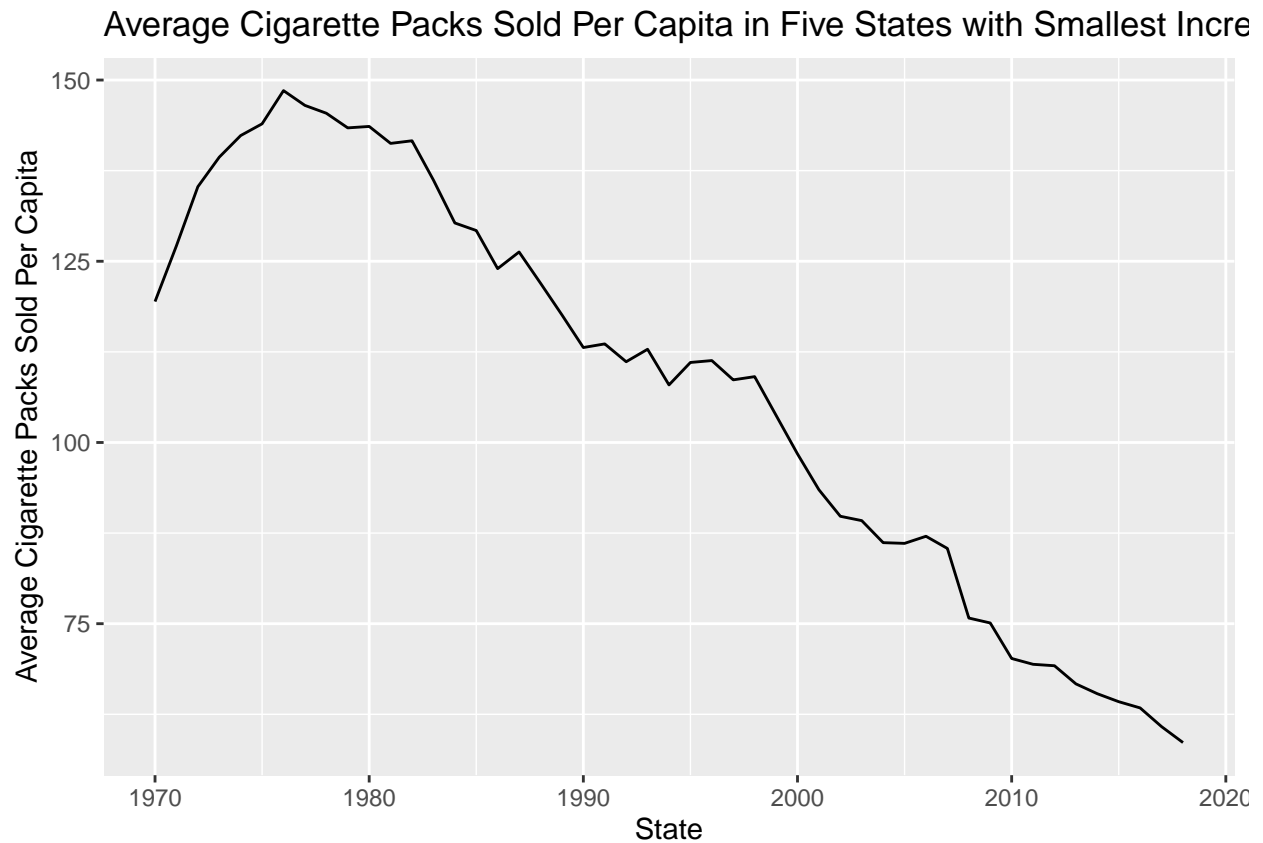
Average Cigarette Packs Sold Per Capita in Five States with Largest Increa

Question 4

```
## # A tibble: 5 x 2
##   state          price_change
##   <chr>                 <dbl>
## 1 North Carolina         4.88
## 2 Georgia                4.87
## 3 Tennessee              4.86
## 4 North Dakota           4.85
## 5 Missouri               4.59
```

Average Cigarette Packs Sold Per Capita in Five States with Smallest Incre

Question 5

Sales from the 5 states with the highest price increase has a steep decline in average cigarette packs sold per capita compared to the 5 states with the lowest price increase. As seen from question 3, there is a sharp decline from 1990 to 2000 whereas the graph from 4 had a more gradual decline.

Question 6

```
##
## Call:
## lm(formula = log(sales_per_capita) ~ log(cost_per_pack), data = subset_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.77629 -0.09967 -0.00787  0.09969  0.78423
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         4.750402   0.008116   585.3   <2e-16 ***
## log(cost_per_pack) -0.171540   0.013829   -12.4   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2107 on 1069 degrees of freedom
## Multiple R-squared:  0.1258, Adjusted R-squared:  0.125
## F-statistic: 153.9 on 1 and 1069 DF,  p-value: < 2.2e-16
```

An 1% increase in the cost of a cigarette pack is estimated to decrease sales by 0.17 percent on average. It is an inelastic relationship.

Question 7

```
##
## Call:
## ivreg(formula = LogSales ~ LogPrices | TotalCigaretteTax, data = subset_data)
##
## Residuals:
##        Min        1Q     Median        3Q       Max
## -0.813864 -0.119209   0.006532   0.117320   0.734440
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.710109   0.009056   520.12   <2e-16 ***
## LogPrices    -0.284348   0.017249   -16.48   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2172 on 1069 degrees of freedom
## Multiple R-Squared: 0.07141, Adjusted R-squared: 0.07054
## Wald test: 271.8 on 1 and 1069 DF,  p-value: < 2.2e-16
```

An 1% increase in the cost of a cigarette pack is estimated to decrease sales by 0.28%. The estimates of those with an instrument are different and shows that a change in cost has a greater impact on sales, making it slightly more elastic. This is because the estimates with an instrument are accounting for the total cigarette tax has on the cost of cigarette packs.

Question 8

```
##     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -1.3090 -0.6156 -0.4092 -0.3572 -0.0524  0.5713


##
## Call:
## lm(formula = log(sales_per_capita) ~ pricehat, data = subset_data2)
##
## Residuals:
##      Min      1Q  Median      3Q     Max
## -0.82897 -0.14423  0.00604  0.14668  1.19203
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.960e-15  9.558e-02    0.00        1
## pricehat    1.000e+00  2.252e-02   44.41   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.28 on 1273 degrees of freedom
## Multiple R-squared:  0.6077, Adjusted R-squared:  0.6074
## F-statistic:  1972 on 1 and 1273 DF,  p-value: < 2.2e-16
```

Question 9

```
##
## Call:
## lm(formula = log(sales_per_capita) ~ log(cost_per_pack), data = subset_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.77629 -0.09967 -0.00787  0.09969  0.78423
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         4.750402   0.008116   585.3   <2e-16 ***
## log(cost_per_pack) -0.171540   0.013829   -12.4   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2107 on 1069 degrees of freedom
## Multiple R-squared:  0.1258, Adjusted R-squared:  0.125
## F-statistic: 153.9 on 1 and 1069 DF,  p-value: < 2.2e-16
```

An 1% increase in the cost of a cigarette pack is estimated to decrease sales per capita by 0.17 percent on average. It is an inelastic relationship.

```
##
## Call:
## ivreg(formula = LogSales ~ LogPrices | TotalCigaretteTax, data = subset_data2)
##
## Residuals:
##        Min        1Q    Median        3Q       Max
## -0.859923 -0.188223 -0.008599  0.202679  1.165439
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.15751    0.02465   209.2   <2e-16 ***
## LogPrices   -0.76265    0.01897   -40.2   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3092 on 1273 degrees of freedom
## Multiple R-Squared: 0.5214,  Adjusted R-squared: 0.521
## Wald test:  1616 on 1 and 1273 DF,  p-value: < 2.2e-16
```

An 1% increase in cost per cigarette pack is estimated to decrease sales per capita by 0.76%. The estimates of those with an instrument are different and show that a change in cost has a greater impact on sales, making it more elastic. This may be due to the fact that the cigarette tax is accounted for and its inflence on the cost per pack.

```
##
## Call:
## lm(formula = log(sales_per_capita) ~ log(tax_dollar), data = subset_data2)
##
## Residuals:
```

```
##      Min      1Q   Median      3Q      Max
## -0.82897 -0.14423  0.00604  0.14668  1.19203
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     4.236866   0.007842  540.26   <2e-16 ***
## log(tax_dollar) -0.480477   0.010820  -44.41   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.28 on 1273 degrees of freedom
## Multiple R-squared:  0.6077, Adjusted R-squared:  0.6074
## F-statistic:  1972 on 1 and 1273 DF,  p-value: < 2.2e-16


##
## Call:
## lm(formula = log(sales_per_capita) ~ pricehat, data = subset_data2)
##
## Residuals:
##      Min      1Q   Median      3Q      Max
## -0.82897 -0.14423  0.00604  0.14668  1.19203
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.960e-15  9.558e-02    0.00        1
## pricehat    1.000e+00  2.252e-02   44.41   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.28 on 1273 degrees of freedom
## Multiple R-squared:  0.6077, Adjusted R-squared:  0.6074
## F-statistic:  1972 on 1 and 1273 DF,  p-value: < 2.2e-16
```

Question 10

Yes, they are different. The elasticity estimates from 1991-2015 are more elastic compared to those from 1970-1990. This may be because taxes on cigarette packs increased and the CDC emphasized the harmful health effects of smoking, making people more sensitive to the change in prices on cigarettes.