Homework 3

Research Methods, Spring 2025

Justin Hahm

My answers to the homework questions are described below. Note that I do the analysis for these answers in a separate `R` script. My analysis file is available in the analysis folder. The GitHub repository for this work is available here. Enjoy!

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. Figure 1

```
Warning: Removed 1 row containing missing values or values outside the scale range
(`geom_bar()`).
```

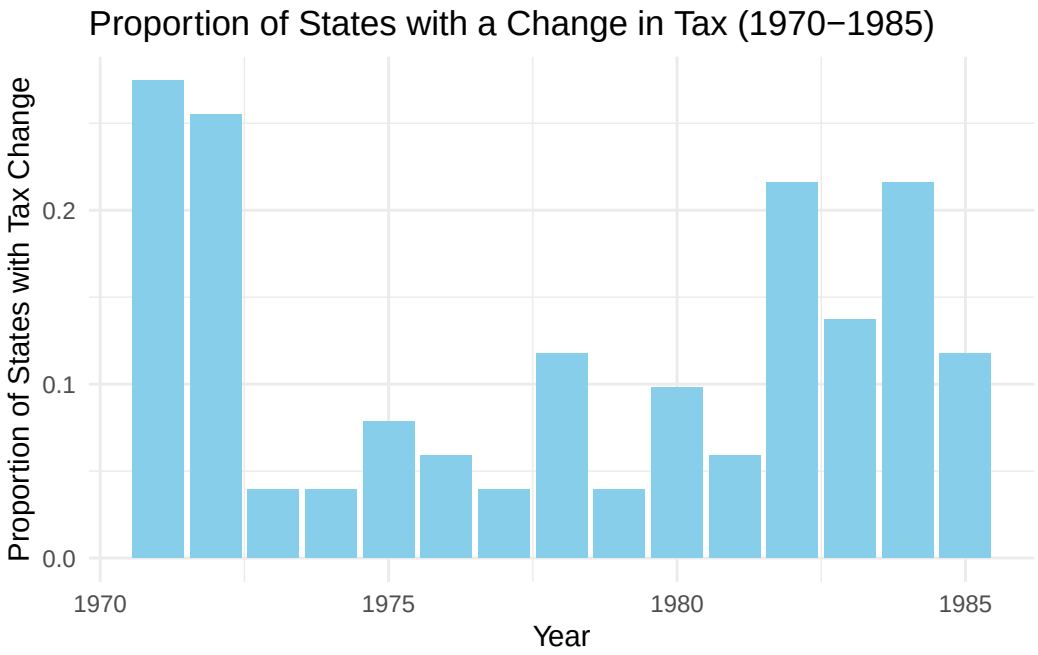

Figure 1

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 2

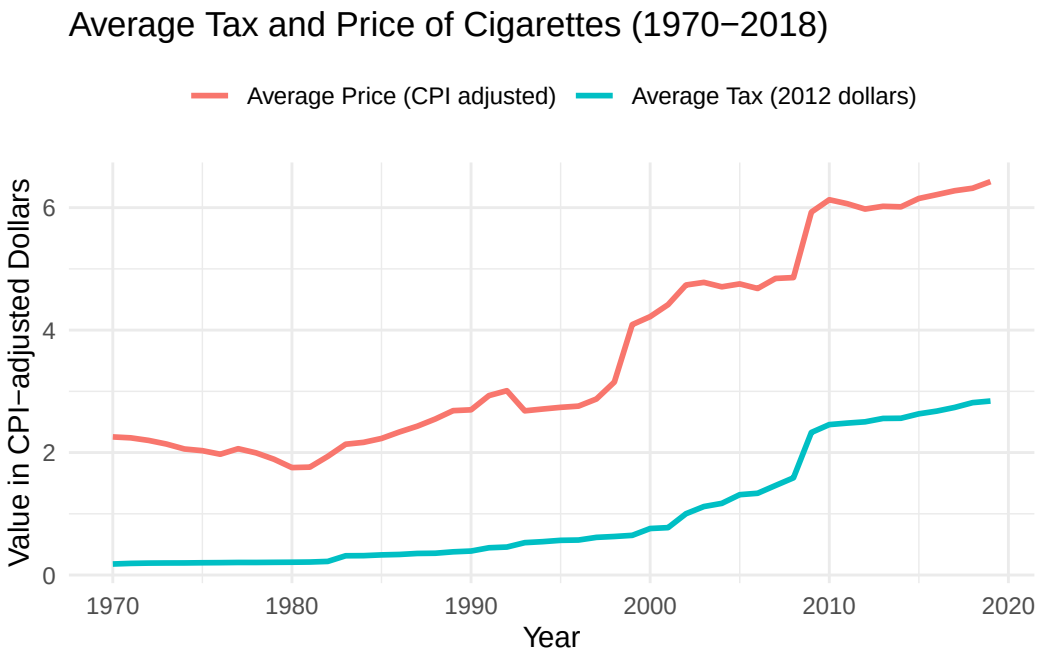

Figure 2

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 3

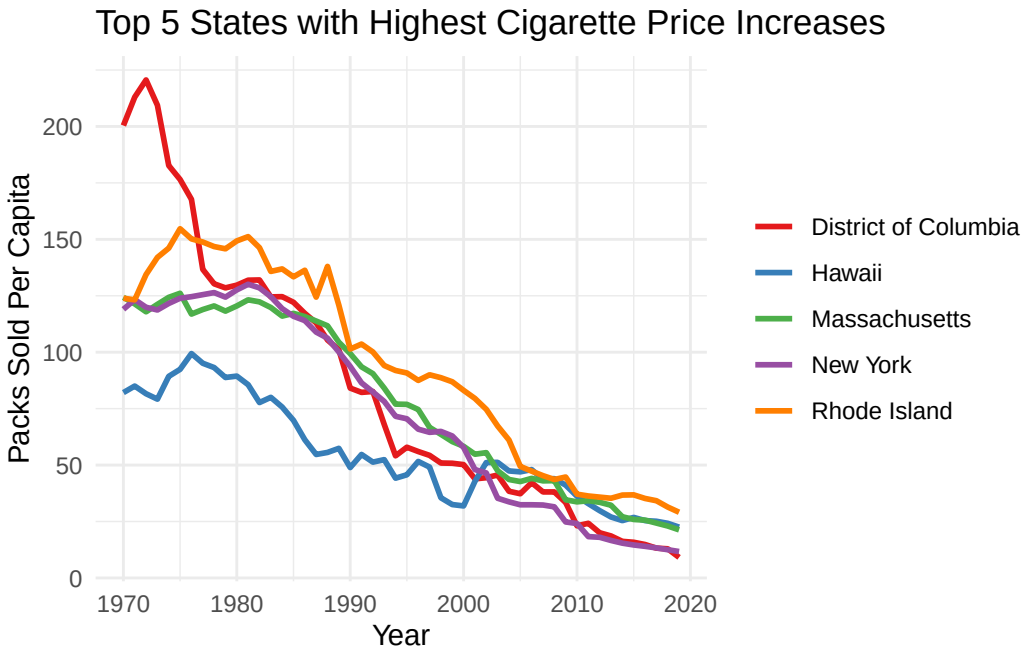

Figure 3

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. Figure 4

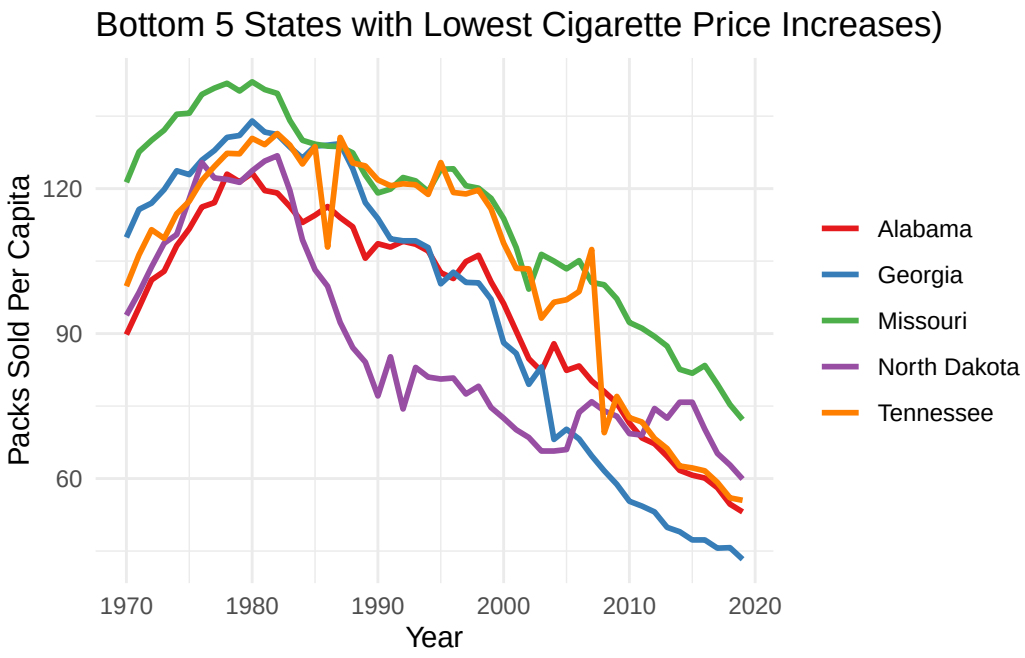

Figure 4

5. Compare the trends in sales from the 5 states with the highest price increases to those with the lowest price increases. Figure 4

The trend in sales is decreasing for both the top 5 and lowest 5 states.

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.

The OLS regression makes it seem that the price elasticity is relatively inelastic.

Table 1: OLS Table

term	estimate	std.error	statistic	p.value
(Intercept)	4.75	0.01	585.32	0
ln_price	-0.17	0.01	-12.40	0

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?

The IV reveals that the coefficient is more elastic.

Table 2: IVS Table

term	estimate	std.error	statistic	p.value
(Intercept)	4.99	0.03	146.34	0
fit_ln_price	0.50	0.09	5.59	0

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

Table 3: Step 1 Table

term	estimate	std.error	statistic	p.value
(Intercept)	-0.48	0.02	-25.3	0
ln_total_tax	-0.41	0.04	-9.4	0

Table 4: Step 2 Table

term	estimate	std.error	statistic	p.value
(Intercept)	4.99	0.02	256.36	0
pricehat	0.50	0.05	9.80	0

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

Table 5: OLS Table

term	estimate	std.error	statistic	p.value
(Intercept)	5.04	0.02	219.93	0
ln_price	-0.67	0.02	-38.09	0

Table 6: Step 1 Table

term	estimate	std.error	statistic	p.value
(Intercept)	1.05	0.01	145.57	0
ln_total_tax	0.73	0.01	64.12	0

Table 7: IVS Table

term	estimate	std.error	statistic	p.value
(Intercept)	5.22	0.03	196.95	0
fit_ln_price	-0.81	0.02	-39.57	0

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

The elasticity estimates are different. The 1991 to 2015 estimates are closer to -1. This is because IV is more closely correlated with the data in this panel. After controlling for the increase in average tax over time (as policies to limit smoking took place) demonstrate that cigarettes became more elastic in recent years. It inutiviely makes sense that culturally cigarettes used to be more inelastic in the past.