

Homework 3

Kendall Pollard

Here is the link to my GitHub Repository: <https://github.com/kpollard8/Homework3>

Here are my answers for Homework 3. I do the coding in a separate R script, but here is the cleaned-up version. I run the analysis separately, save the workspace with only the summary stats, figures, and tables that I need, and then load the workspace in the final qmd. My analysis file with answers and code to all the questions is available in the analysis folder.

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

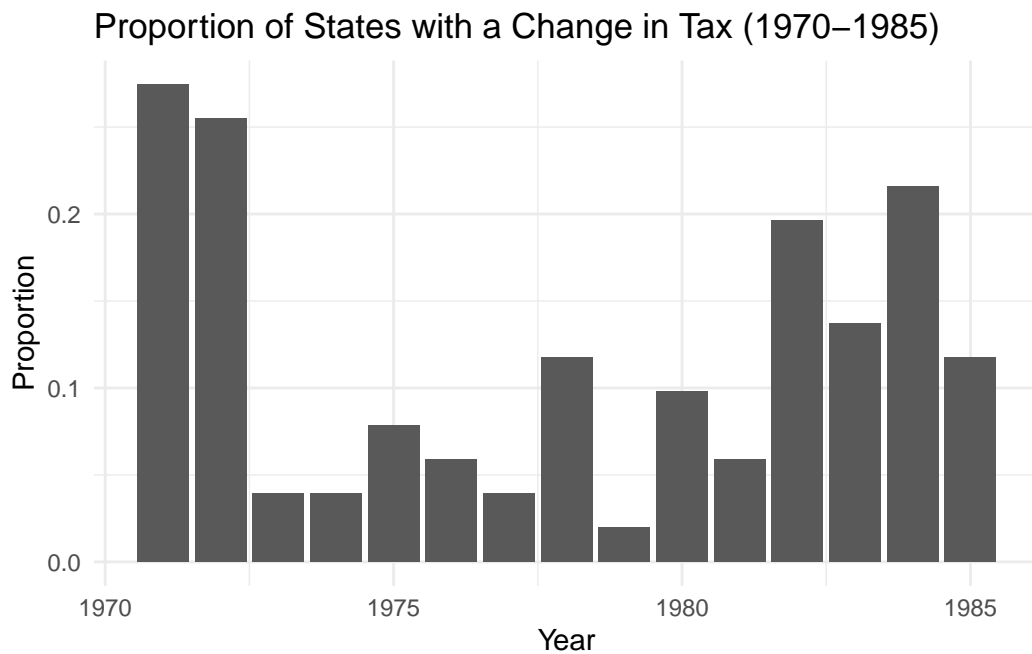


Figure 1: Question 1 Graph

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.

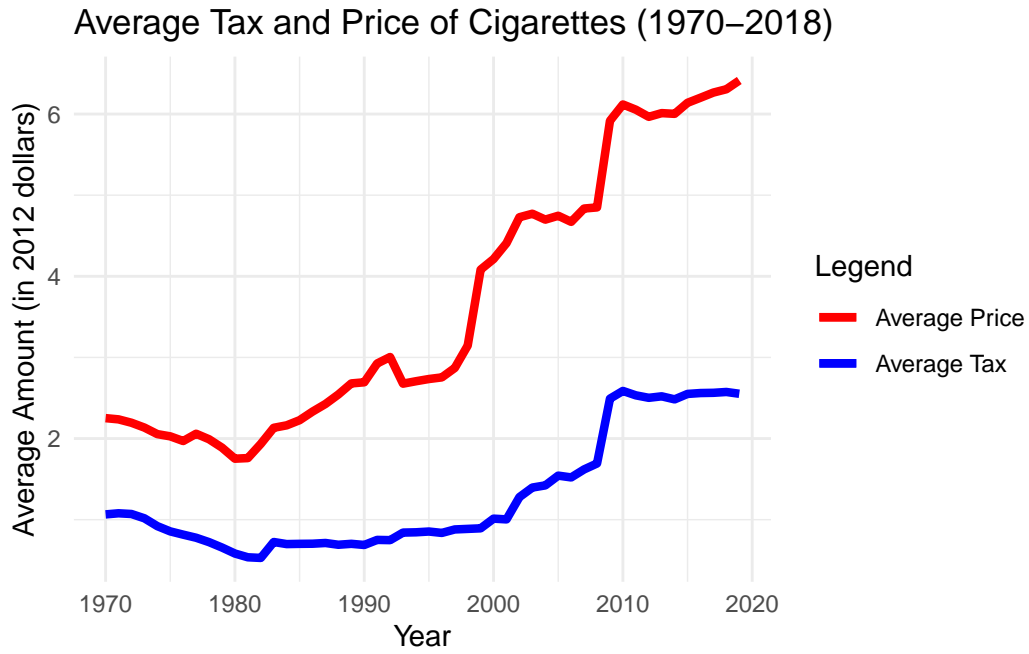


Figure 2: Question 2 Graph

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.

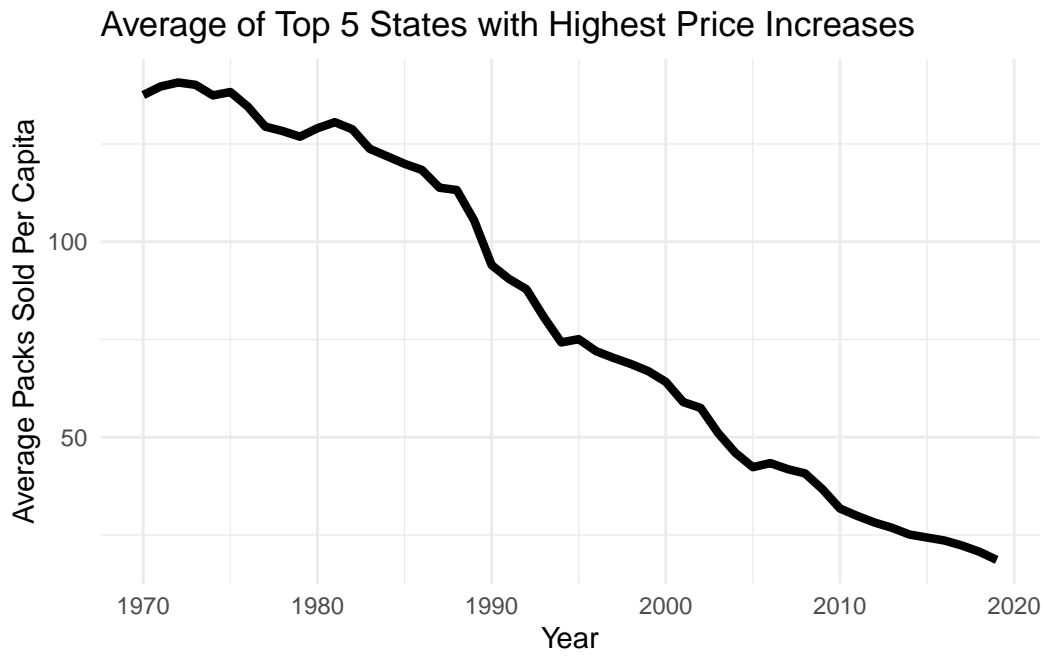


Figure 3: Question 3 Graph

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.

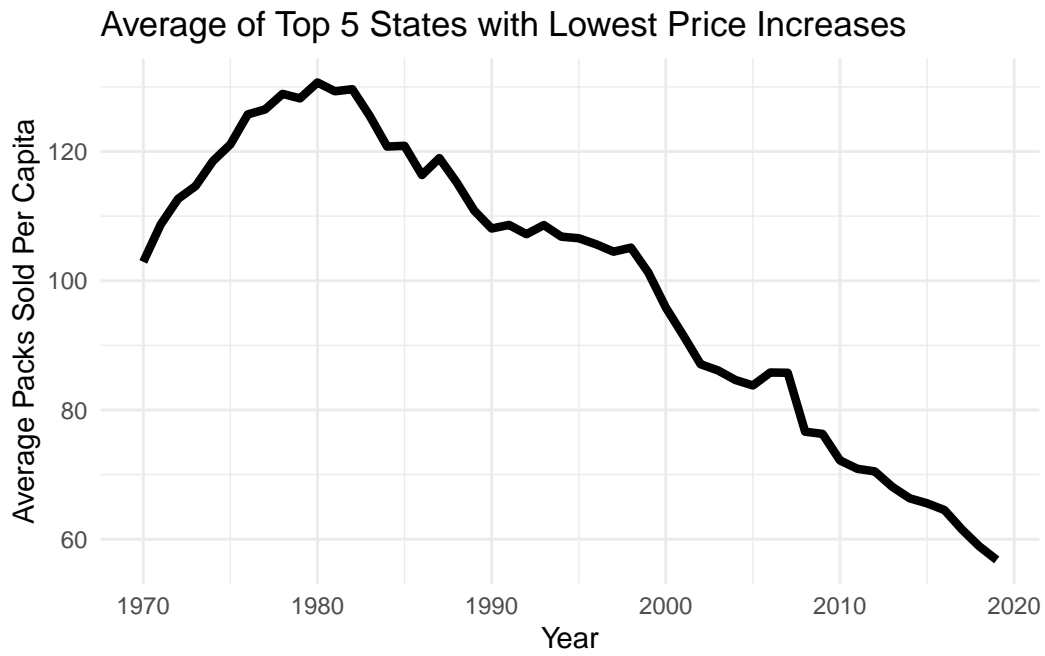


Figure 4: Question 4 Graph

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

Both graphs are overall decreasing average packs sold from 1970 to 2020. However, there is a clear difference between the two from 1970 to 1980. In the states with the lowest price increases, sales were still increasing from 1970 to 1980. After 1980, sales started to decline, but at a slower rate than the 5 states with the highest price increase. This makes sense with common laws of demand and supply; demand will decrease with price increases, but at a higher rate with higher price increases.

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.

regression_table

Table 1: Log Regression

Betas	Coefficients	Standard Errors
Intercept	5.4273812	0.0297515
Log Price	-0.8094384	0.0383656

The intercept (5.43) represents the estimated log sales when log prices are zero. This value is highly significant (p value very small) indicating a strong positive relationship between log prices and log sales. The coefficient for log prices (-0.8094384) represents the estimated change in log sales for a one percent increase in log prices. A one percent increase in price decreases sales by 0.809%. This coefficient is also highly significant, with another very small p value, suggesting a negative relationship between log prices and log sales.

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?

Table 2: IV Model for Question 7

Betas	Coefficient	Standard Error
Intercept	5.4167966	0.0544940
Log Price	-0.7955235	0.0712351

The estimated intercept is 5.42, which represents the estimated log sales when log prices are zero. This value is highly significant, with a p value close to 0, indicating a strong positive relationship between log prices and log sales when using the instrument.

Coefficient for Log Prices (Price Elasticity): The estimated coefficient for log prices is -0.7955235. This coefficient represents the estimated change in log sales for a one percent increase in log prices, considering the influence of the instrument. This is again intuitive because as price increases, sales would decrease, in this case by 0.79%. It is highly significant (small p value), suggesting a negative relationship between log prices and log sales, indicating that as log prices increase, log sales tend to decrease.

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

First stage and reduced form

Table 3: First and Second Stages for Question 8

Regression	Coefficient	Standard Error
First Stage		
Intercept	0.839645686322117	0.00542130797936402
Log Tax	0.260060067049418	0.012442534888896
Reduced Form		
Intercept	4.74883874500992	0.00920153912987099
Log Tax	-0.206883898032473	0.0211186068179054

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

Question 6 repeated for 1991-2015

Table 4: Log Regression 2

Betas	Coefficients	Standard Errors
Intercept	5.6599553	0.0363844
Log Price	-0.9968136	0.0246921

The coefficient for log prices (-0.9968136) represents the estimated change in log sales for a one percent increase in log prices. A one percent increase in price decreases sales by 0.99%. It is a nearly one to one relationship in this time period. This coefficient is also highly significant, with another very small p value.

Question 7 repeated for 1991-2015

Table 5: IV Model for Question 7

Betas	Coefficient	Standard Error
Intercept	5.879862	0.0407797
Log Price	-1.150084	0.0278109

The estimated coefficient for log prices is -1.150084. This coefficient represents the estimated change in log sales for a one percent increase in log prices, considering the influence of the instrument. So, sales decrease by -1.15% when price increases by 1%. So, this is an elastic good.

Question 8 repeated for 1991-2015

First stage and Reduced Form

Table 6: First and Second Stages for Question 8 repeated

Regression	Coefficient	Standard Error
First Stage		
Intercept	1.31507299550809	0.00438589379991311
Log Tax	0.513550303745968	0.00692246770990574
Reduced Form		
Intercept	4.36741790115986	0.00843951714008135
Log Tax	-0.590625849539619	0.0133204969282582

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

The coefficient for log prices remains consistently strongly negative across both time periods. It becomes even more negative in 1991-2015, by about 0.19 for the elasticity model, 0.3 for the IV model. So, as time went on, the coefficients got even more negative, proving that the good is becoming more elastic over time.

The instrument is much stronger in the second period, which makes sense because that is when the biggest tax increases happened.

The intercept values also remain consistent between the two periods, with estimated log sales of between 4 and 5 when log prices are zero for the elasticity and IV models. These intercept values are highly significant in both regressions, indicating a strong positive relationship between log prices and log sales. For the first stage and reduced form, the intercepts are both around 1 in each time period.

This analysis is intuitive because more recently, we have become aware of the health issues with cigarette smoking. It makes sense that the good is becoming more elastic because we are more aware that cigarettes are bad so we are becoming more price-sensitive.