Homework 3

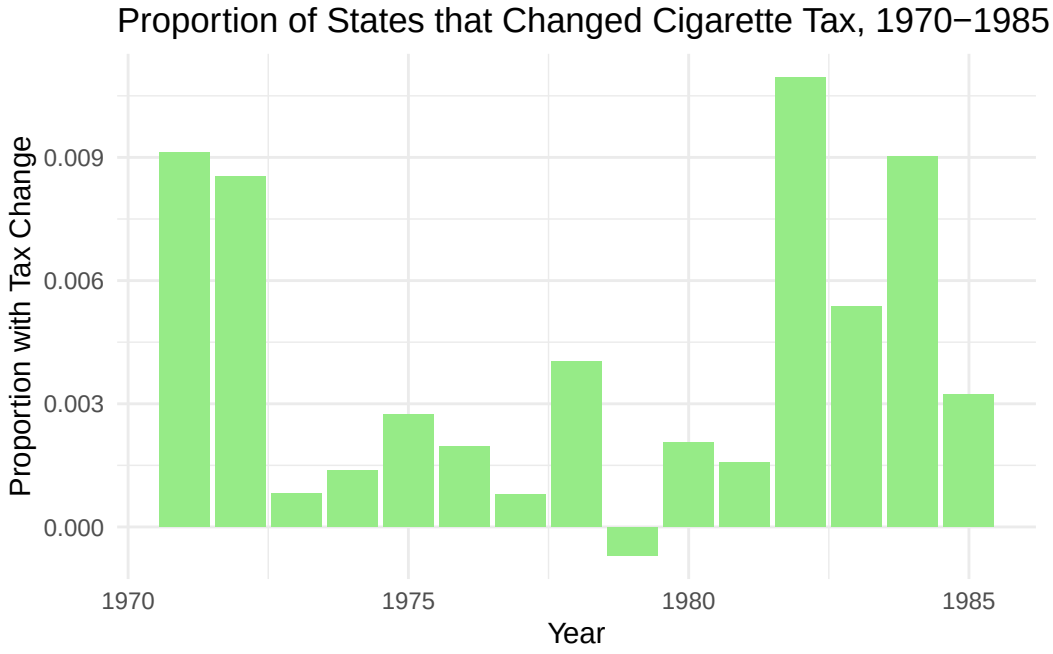
Research in Health Economics, Spring 2025

Sammy Ramacher

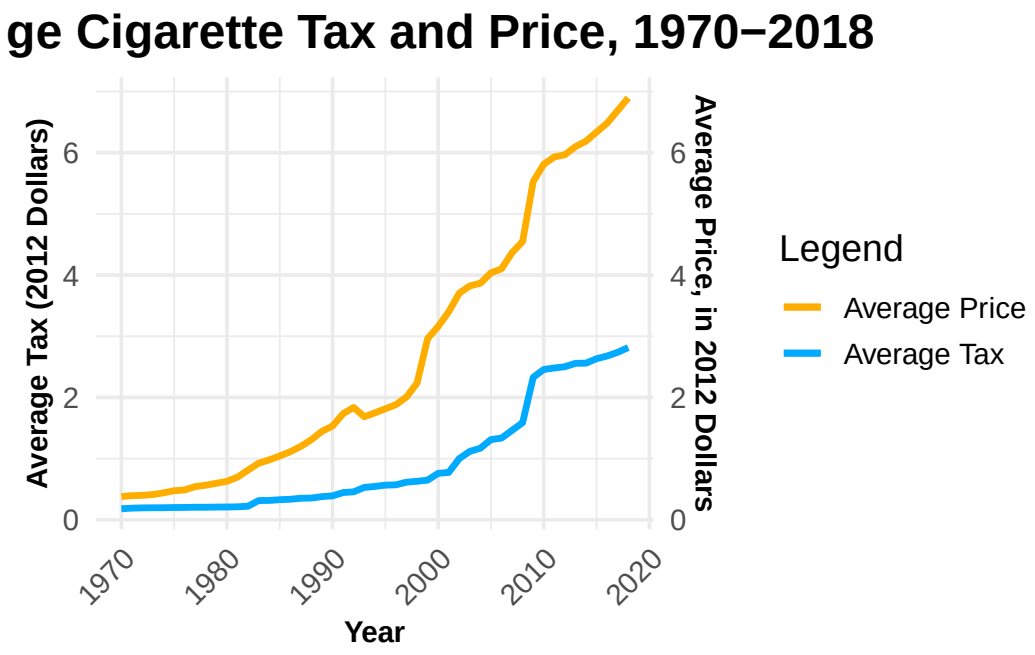
The GitHub repository for this work is available here.

Summarize the Data

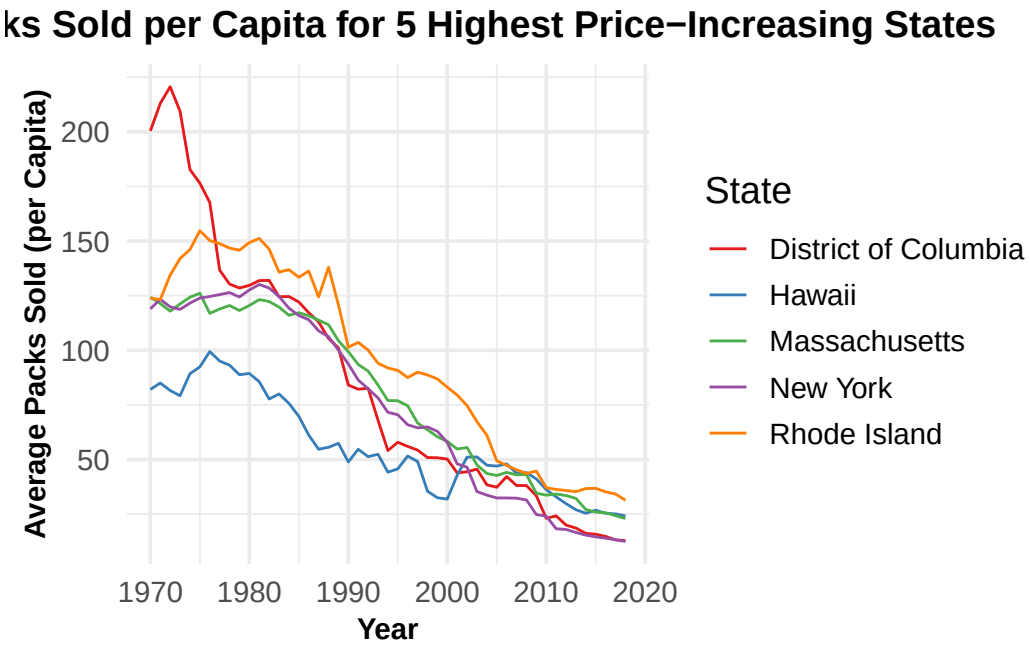
Question 1. Present a bar graph showing the proportion of states with a change in their cigarette tax in each year from 1970 to 1985.



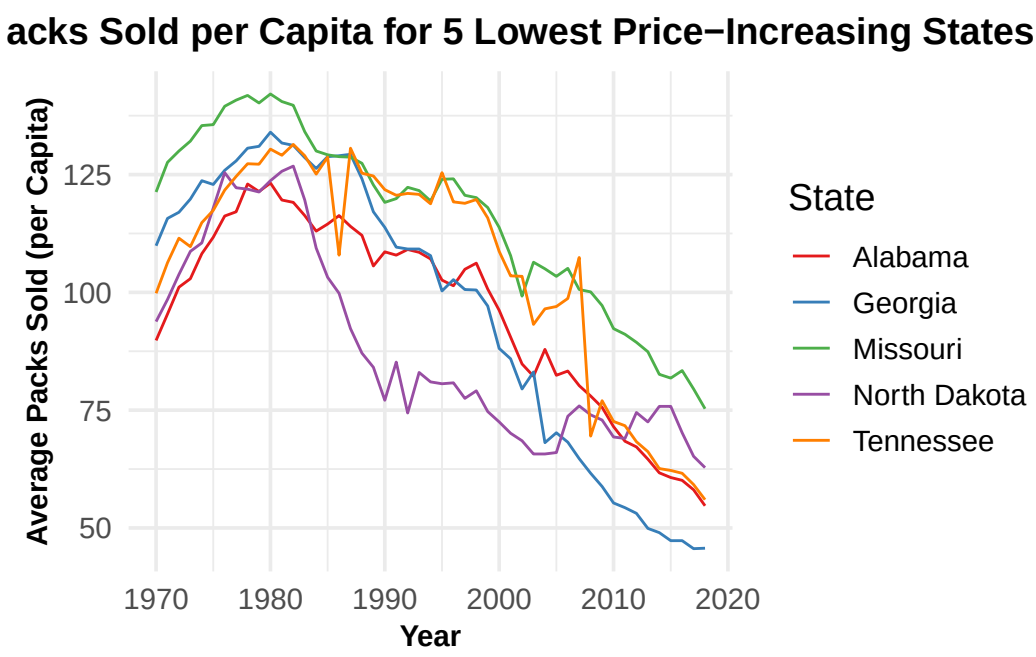
Question 2. Plot on a single graph the average tax (in 2012 dollars) on cigarettes and the average price of a pack of cigarettes from 1970 to 2018.



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 5 states with the highest increase in cigarette prices have a sharp decline over time in the average number of packs sold. Meanwhile, the 5 states with the lowest price increase for cigarettes have a more stable average number of packs sold over time, although there is still a steady declune. Both datasets have a maximum average number of packs sold around 1980, and then the averages drop over time. All 5 of the states with the highest price increase sell less packs per capita than all 5 of the low price increase states in later years, starting around 2015.

Estimate ATEs

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.

Table 1: Log Sales on Log Price Regression Results

term	estimate	std.error	statistic	p.value
(Intercept)	4.750	0.008	585.32	0
ln_price_2012	-0.172	0.014	-12.40	0

These results indicate that a 1% increase in cigarette price leads to a 0.172% decrease in sales. Because the absolute value of the coefficient for price is less than 1, the demand for cigarettes is relatively inelastic.

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?

Table 2: Log Sales on Log Price Regression with Instrument Results

term	estimate	std.error	statistic	p.value
(Intercept)	5.158	0.025	209.25	0
fit_ln_price_2012	-0.763	0.019	-40.20	0

These results indicate that a 1% increase in price leads to a 0.763% decrease in sales. This is different than the first regression that does not have the instrument. The price elasticity for the regression with the instrument is much larger, making the demand for cigarettes more elastic when controlling for endogeneity. Therefore, the first model is likely underestimating the true price elasticity of demand.

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

Table 3: First Stage Results from Instrument

term	estimate	std.error	statistic	p.value
(Intercept)	1.207	0.005	242.91	0
ln_tax_2012	0.630	0.007	91.88	0

Table 4: Reduced-Form Results from Instrument

term	estimate	std.error	statistic	p.value
(Intercept)	4.375	0.025	176.63	0
ln_tax_2012	-0.307	0.017	-18.17	0

Question 9. Repeat questions 6-8 focusing on the period from 1991 to 2015.

Table 5: Log Sales on Log Price Regression Results: 1991-2015

term	estimate	std.error	statistic	p.value
(Intercept)	5.039	0.023	219.93	0
ln_price_2012	-0.666	0.017	-38.09	0

Table 6: Log Sales on Log Price Regression with Instrument: 1991-2015

term	estimate	std.error	statistic	p.value
(Intercept)	5.158	0.025	209.25	0
fit_ln_price_2012	-0.763	0.019	-40.20	0

Table 7: First Stage Results from Instrument: 1991-2015

term	estimate	std.error	statistic	p.value
(Intercept)	1.207	0.005	242.91	0
ln_tax_2012	0.630	0.007	91.88	0

Table 8: Reduced Form Results from Instrument: 1991-2015

term	estimate	std.error	statistic	p.value
(Intercept)	4.237	0.008	540.26	0
ln_tax_2012	-0.480	0.011	-44.41	0

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

The elasticity estimators for 1991-2015 between the regular logistic regression and the logistic regression with the estimator are closer to each other than for the years 1970-1990. This means that the endogeneity controlled for by the tax instrument has less of an effect in the years 1991-2015. The elasticity of demand of the logistic regressions with the instruments in both sets of years are the same; they are both -0.763. This means that the effect of price on ciragrette sales are the same in both time periods.