A Zero inflated Negative Binomial Model for Cigarette consumption in the U.S Post MSA

Peter Amato, Mohammad Iqbal, Anthony Stefanidis

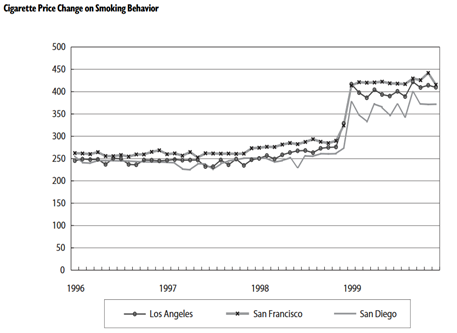
Non-Linear Econometrics (Eco 722)

Professor Colman

Introduction

Mei-Ling author of “*The effect of a major cigarette price change on smoking Behavior in California: a zero-inflated negative binomial model”* was instrumental in supporting the claim that an increase in the price of cigarettes affects the consumption of cigarettes. The author used a ZINB (zero inflated negative binomial) model to examine the relationship between prices of cigarettes and the consumption of cigarettes. She collected her sample from the 1996-1999 Behavioral Risk Factor Survey (BRFS). The Bureau of Labor Statistics also collected the CPI of Tobacco products at the county level prior to 2004. We on the other hand used the country wide Tobacco CPI in lieu of the county data. California was of great interest to researchers studying Tobacco use because they increased the tax on tobacco three times since 1989. Proposition 99 was a law that increased the tax on cigarettes by 250% (paper) in 1989. The last tax increase in our period of examination was proposition 10, which increased the tax on a pack of cigarettes by 43 %( paper). It is important we look at price increase in general because non-tax oriented effects increased the price of cigarettes in that period. The most notable of which is the Master Settlement Agreement (MSA) in 1998, which was such a large lawsuit against nationwide cigarettes manufacturers that it forced them to raise the average price of a pack by 40 cents to compensate for lost revenue and settlement payments.

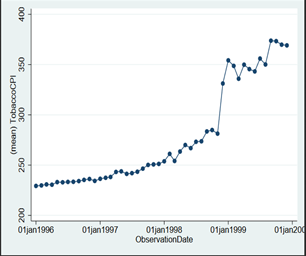
Mei Lings Cigarette Price Change on Smoking Behavior



Tobacco Products CPI V.S Year for Three Counties of Interest

Mei’s graph shows a slight increase or no change in Tobacco CPI relative to the base level of tobacco products, and then an explosion to the CPI in mid-1998. This presents the perfect opportunity to study the effects of a natural experiment, as the MSA clearly drove up the average price of cigarettes.

Cigarette Price Change on Smoking Behavior



National Tobacco Products CPI V.S Year

Our study of the incremental tax increases on cigarettes and other tobacco products shows that the CPI grew more nationally over the years than in the California counties previously listed. A regression of cigarette price on year shows a stronger positive effect, than the last graph depicts. Both graphs depict the effects of the MSA and provide strong evidence for the rationale that the MSA’s effect on cigarette prices was strong and deliberate. This provides a great context for the study of price elasticity of cigarettes, and how that price hike effects consumption. This was the main focus of Mei’s paper and in our research.

Literature Review

One of the main purposes of Mei-Ling’s paper is to reaffirm the results of previous papers. The price elasticity of smoking has long been a subject of study. Sung et al (paper). Used panel data from 11 western states from 1969 -1990, and found that the price elasticity was -0.40 in the short run and -0.48 over the long run. Keeler et al (paper) studied cigarette demand using time-series data from 1980 – 1990 in California, finding that the short-run price elasticities were between -0.30 & -0.50, while the long run elasticities were between -0.50 & -0.60. Hu et al later analyzed the impact of proposition 99 and the accompanying anti-smoking campaign using the same survey data for 1985 to 1991. They found that the price elasticity of cigarettes increased to -0.30.

Data

Our paper examines the same effects of cigarette price elasticity in California using the responses to the BRFS survey and Tobacco CPI data for the whole country. The original author used the California specific CPI for tobacco products to examine elasticities, but this data is no longer available to the public. We will use a two- stage model to measure cigarette consumption (dependent variable), dependent on the natural log of the tobacco CPI, using the same approach as Mei.

The appended dataset for San Diego, Los Angeles, and San Francisco from the BRFS 1996 – 1999 surveys contains 14,235 observations, while the original paper has 16,260 observations. The author’s final sample contains 11,180 observations after dropping regression variables whose values are missing, while our sample drops from 14,235 to 13,435 values. These datasets were also created by dropping everybody who does not live in Los Angeles, San Diego, and San Francisco County. A large source of error may be the discrepancy between our sample sizes. We could not figure out what caused the small difference in size even though we followed the data cleaning procedures outlined in the paper. Our averages are not substantially different though. Our dependent variable is number of cigarettes smoked per day.

The dependent variable of our first stage logistic is issmoker, which designates whether or not the person is a smoker (0 if they are not, 1 if they are). The dependent variable for our second stage negative binomial is lncigaretteprice, the natural log of the CPI for tobacco products. We also generated a variable for isnonsmoker, which takes on the opposite value. Lncigaretteprice is our primary dependent variable, and the main variable of interest in our paper. We took the natural log of tobacco CPI to normalize the tobacco CPI, so we could interpret model changes as percentages. We then generated several categorical variables which take on the value 1 if true, and 0 if false.

We generated ethnicityblack, which designates whether or not a person is black. We generated ethnicityhispanic, which designates whether or not a person is hispanic. We generated ageisbetween25and34, which designates whether or not the person is between the age of 25 and 34. We generated ageisbetween35and44, which designates whether or not the person is between the age of 35 and 44. We generated ageisbetween45and54, which designates whether or not the person is between the age of 45 and 54. We generated ageover65, which designates whether or not the person is over the age of 65. We generated ismale, which designates whether or not the person is a male. We generated issingle, which designates whether or not the person is single. We generated isdivorced, which designates whether or not the person is divorced. We generated iswidowed, which designates whether or not the person is widowed. We generated hastakensomecollegeclasses, which designates whether or not the person has taken some college classes (is a college student). We generated collegegraduate, which designates whether or not the person is a college graduate. We generated isunemployed, which designates whether or not the person is an unemployed. We generated incomebetween10kand20k, which designates whether or not the person has an annual income between $10,000 and $19,999. We generated incomebetween20kand35k, which designates whether or not the person has an annual income between $20,000 and $34,999. We generated incomeover35k, which designates whether or not the person has an annual income of $35,000 or more. We generated isincomeunknown, which designates whether or not the person has an annual income that is unknown or was not filled out on the survey. We generated isgenhlthgood, which designates whether or not the person is in good health or not. We generated isgenhlthfair, which designates whether or not the person is in fair health or not. The variable iyear is a variable generated by the BRFSS, which is a variable that indicates which year the survey was administered in.

The discrepancy in sample size compelled us to study our sample further. The following diagrams depict the characteristics of Mei’s sample and our sample.

Mei’s Sample Summary

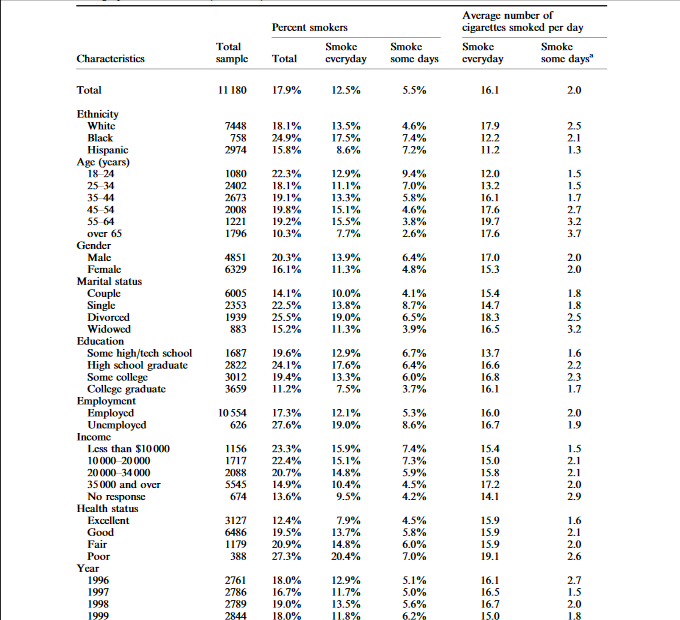


Table 1 shows that 17.9% of the individuals in the sample were current smokers with 12.5% regular smokers and 5.5% irregular smokers. Mei constructed the number of cigarettes smoked per day for irregular smokers by multiplying their number of days of smoking by the number of cigarettes smoked, and then divided by 30 days to create a monthly average. Table 1 indicates that the average number of cigarettes consumed per day between regular and irregular smokers is quite different, 16.1 and 2.0, respectively. Blacks had the highest smoking prevalence, 24.9%, while Whites consumed the highest number of cigarettes, 17.9 per day, among those regular smokers. Those aged 18-24 had the highest smoking prevalence, but the older age brackets had higher average cigarette consumption. Males had higher smoking prevalence than females. Couples had the lowest smoking prevalence. High school graduates had the highest smoking prevalence while college graduates had the lowest. The unemployed had a higher smoking prevalence rate than the employed. Individuals who had income $35 000 or above had the highest smoking prevalence and average cigarette consumption, while those with unknown income had the highest average cigarette consumption for irregular smoking. People with poor health status had the highest smoking prevalence, and saw a large jump in average cigarette consumption.

Our Sample Summary

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Percent Smokers | Percent Smokers | Percent Smokers | Average Number Of Cigarettes Smoked Per Day | Average Number Of Cigarettes Smoked Per Day |
| Characteristics | Total Sample | Total | Smoke Everyday | Smoke Some Days | Smoke Everyday | Smoke Some Days |
| Total | 13,545 | 21.54% | 16.29% | 5.25% | 18.29 | 0.70 |
| Ethnicity |  |  |  |  |  |  |
| White | 9,875 | 22.3% | 17.85% | 4.45% | 19.61 | 0.70 |
| Black | 1,616 | 22.03% | 15.41% | 6.62% | 14.45 | 0.91 |
| Hispanic | 2,054 | 22.28% | 12.07% | 10.21% | 12.39 | 0.59 |
| Age(years) |  |  |  |  |  |  |
| 18-24 | 1,349 | 23.87% | 16.53% | 7.34% | 14.38 | 0.38 |
| 25-34 | 2,916 | 22.97% | 16.15% | 6.82% | 16.09 | 0.63 |
| 35-44 | 3,162 | 25.01% | 19.73% | 5.28% | 18.55 | .77 |
| 45-54 | 2,414 | 24.23% | 19.55% | 4.68% | 20.21 | 0.70 |
| 55-64 | 1,494 | 20.82% | 16.27% | 4.55% | 20.79 | 0.83 |
| Over 65 | 2,176 | 10.80% | 7.86% | 2.94% | 20.63 | 1.18 |
| Gender |  |  |  |  |  |  |
| Male | 5,699 | 24.06% | 18.20% | 5.86% | 19.93 | .77 |
| Female | 7,846 | 19.71% | 14.91% | 4.80% | 17.03 | 0.65 |
| Marital Status |  |  |  |  |  |  |
| Couple | 7,167 | 16.25% | 12.68% | 3.57% | 18.36 | 0.75 |
| Single | 2,759 | 24.93% | 17.14% | 7.79% | 16.27 | 0.65 |
| Divorced | 1,937 | 31.96% | 25.35% | 6.61% | 20.64 | 0.61 |
| Widowed | 1,229 | 16.11% | 12.29% | 3.82% | 18.52 | 0.93 |
| Education |  |  |  |  |  |  |
| Some High/tech School | 1,095 | 31.32% | 24.11% | 7.21% | 18.78 | 0.85 |
| High School Graduate | 3,816 | 26.8% | 21.43% | 5.37% | 18.52 | 0.63 |
| Some College | 3,888 | 23.58% | 17.74% | 5.84% | 18.62 | 0.80 |
| College Graduate | 3,938 | 13.39% | 9.17% | 4.22% | 17.44 | 0.59 |
| Employment |  |  |  |  |  |  |
| Employed | 8,449 | 23.19% | 17.45% | 5.74% | 17.79 | 0.65 |
| Unemployed | 564 | 32.98% | 24.47% | 8.51% | 19.16 | 0.63 |
| Income |  |  |  |  |  |  |
| Less than $10,000 | 1,153 | 24.29% | 17.35% | 6.94% | 17.32 | 0.85 |
| $10,000 - $20,000 | 2,013 | 26.18% | 20.17% | 6.01% | 17.84 | 0.54 |
| $20,000-$34,000 | 2,995 | 25.34% | 19.90% | 5.44% | 17.77 | 0.70 |
| $35,000 And Over | 5,930 | 18.11% | 13.42% | 4.69% | 18.74 | 0.66 |
| No Response | 697 | 20.52% | 14.78% | 5.74% | 19.58 | 0.85 |
| Health Status |  |  |  |  |  |  |
| Excellent | 3,305 | 15.64% | 10.92% | 4.72% | 16.85 | 0.61 |
| Good | 3,676 | 24.87% | 19.03% | 5.84% | 19.23 | 0.63 |
| Fair | 1,510 | 9.62% | 4.72% | 4.90% | 18.53 | 1.00 |
| Poor | 538 | 9.48% | 2.42% | 7.06% | 22.2 | 1.20 |
| Year |  |  |  |  |  |  |
| 1996 | 3,203 | 22.51% | 17.39% | 5.12% | 17.60 | 0.73 |
| 1997 | 3,253 | 21.15% | 15.92% | 5.23% | 19.25 | 0.74 |
| 1998 | 3,348 | 22.19% | 17.23% | 4.96% | 18.56 | 0.67 |
| 1999 | 3,741 | 20.48% | 14.84% | 5.64% | 18.21 | 0.69 |

Our Table shows that 21.54% of the individuals in the sample were current smokers with 16.29% regular smokers and 5.25% irregular smokers. We constructed the number of cigarettes smoked per day for irregular smokers by multiplying their number of days of smoking by the number of cigarettes smoked, and then divided by 30 days to create a monthly average. The daily cigarette consumption for daily smokers was only 2.19 higher on average, but the average number of cigarettes smoked is 1.3 lower. Mei’s average non-daily smoker cigarette consumption is one standard deviation above ours. Whites had the highest smoking prevalence, 22.01%, but blacks were only 0.27 percentage points behind. Whites still consumed far more cigarettes everyday than blacks, just like Mei’s sample. Those aged 35-44 had highest smoking prevalence, but our sample had lower smoking prevalence as age increased, and increased average number of cigarettes smoked, just like Mei’s sample. Males had higher smoking prevalence than females. Couples had the lowest smoking prevalence. People who were enrolled in some High School or Tech School had the highest smoking prevalence while college graduates had the lowest. The unemployed had a higher smoking prevalence rate than the employed. Individuals who had an income of $35,000 and over had the lowest smoking prevalence in our sample, but average cigarette consumption in both groups was not very different. People with good health status had the highest smoking prevalence. This is a huge problem considering Mei’s sample logically had people with poor health as the most likely candidate to smoke, but we theorize people with poor health status were forced to quit smoking, or are no longer even able to smoke. We also have very few individuals in poor health, so small movements in variance can cause huge percentage point shifts in proportion. We found the same trend as Mei in cigarette consumption by health status, in that the average number of cigarettes smoked in both groups increased as health status decorates.

Statistical Methods and Regression

The baseline group in the regression analysis is composed of those who are white, age 18–24, female, couples, with some high school (or technical school) education, employed, with income less than $10 000, and with excellent health condition. There is definitely a problem of excess zeros, since only 21.54% of the individuals in our sample are smokers. We still wanted to run an equidispersion test and summarize. We used a logistic model to select for smokers, and a negative binomial to measure cigarette consumption. We did this because Mei believed it was the best method given the ZINB method, where:

Yi = with probability qi

Yi = Poisson (λi)

Where qi = (exiY) (1 + exiY)

And Yi = Zi\*Y\*I or smokenum = (issmoker)\*(latent unobservable desire to smoke)

With µ = E (yi) = qi0 + (1 – qi) λi = (1 – qi) λi & Var (qi) = λi (1 – qi) (1 + λiqi)

Var (yi)/E (yi) will approach a Poisson distribution as qi approaches 0, which is handy for our sample since 78.46% of our sample has a zero value for Yi.

We approach a NB distribution, so we have a Zero inflated Binomial Model with variance:

Var (yi) = λi \* (1 – qi) \* [1 + λi \* (qi + α)]

The b1 on lncigaretteprice is an elasticity assuming b1xi is equal to b1 (lncigaretteprice) where lncigaretteprice is a regressor in xb.

lny = ln(xb)

ε = ΔQ/ΔP

dlny/dy = 1/y

dlny = dy/y = xy/y

ε = dlny/dlnx

Results

ZINB Alpha Test

/lnalpha | .0123497 .0294985 0.42 0.675 -.0454664 .0701657

---------------------------+----------------------------------------------------------------

Alpha | 1.012426 .0298651 .9555517 1.072686

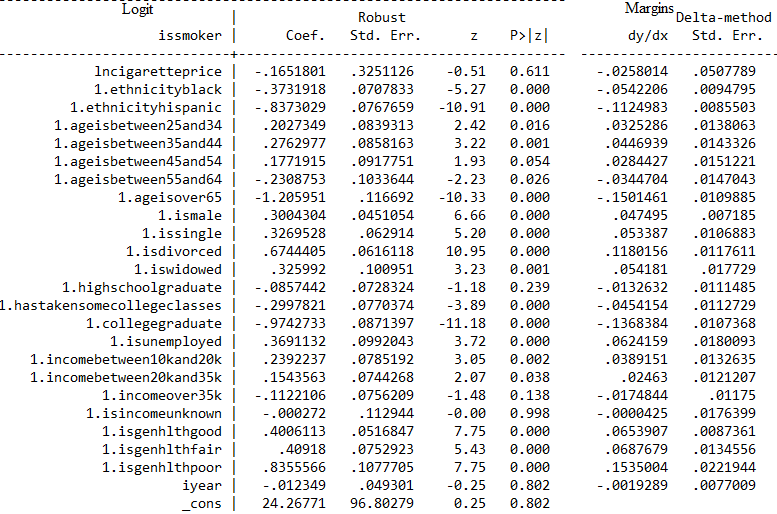
--------------------------------------------------------------------------------------------

Likelihood-ratio test of alpha=0: chibar2 (01) = 1.9e+04 Pr>=chibar2 = 0.0000

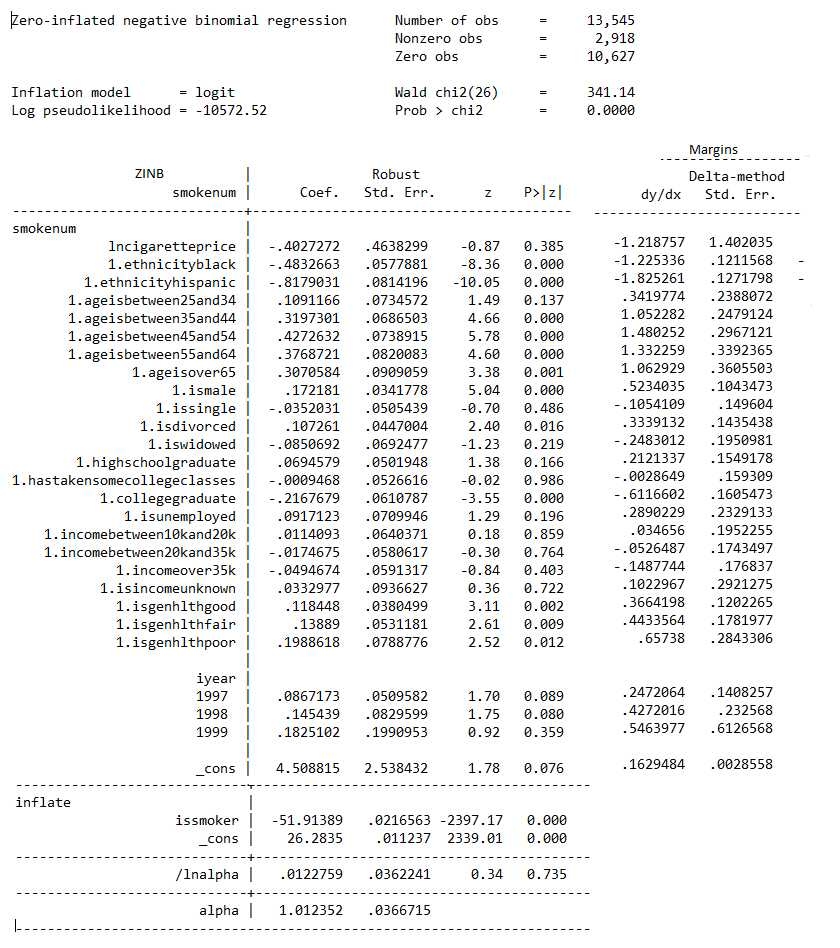
Vuong test of zinb vs. standard negative binomial: z = 90.32 Pr>z = 0.0000

I ran a ZINB to test for equidispersion. A test for equidispersion against the chi squared distribution where the null is equidispersion produces a p-value of 0.00. This is expected considering the negative zero inflated binomial regression (ZINB) produced an alpha of 1.01 and Ln alpha of 0.01. A simple summary of smokenum produces a mean of 3.03 and standard deviation of 8.22. We are clearly over dispersed. A Vuong test of our ZINB model produced a z-score = 90.32 with a p-value of 0.000, demonstrating that the ZINB model is preferable to the negative binomial model. Mei did not report the margins for her regression, but instead summarized it in her findings as trends. We will go into more detail in order to see the difference between the participation and intensity effect.

Margins for Original Model



It’s important to note Mei used a logit as a selection model and participation equation for her ZINB model, modelling only nonsmoker outcomes as a positive outcome. We used a logit for issmoker as our selection model and participation. A marginal analysis of a logistic model of whether or not a person is a smoker reveals a lot about participation effects. Your odds **relative to the baseline group** of being a smoker decrease by .258 percentage points, or PP, for every 10% increase in the price of cigarettes. Your odds being a smoker decrease by 5.42 PP if you are black, and decrease by 11.25 PP if you are Hispanic. Your odds of being a smoker increase by 3.25 PP if your age is between 25 and 34, 4.47 PP if your age is between 35 and 44, 2.84 PP if your age is between 45 and 54, and decreases by 3.44 PP if your age is between 55 and 64, or 15.01 PP if you are over 65 years old. Compared to females your odds of being a smoker increase by 4.75 PP if you are male. Your odds of being a smoker increase by 5.34 PP if you are single, 11.8 PP if you are divorced, and 5.42 if you are widowed. Your odds of being a smoker decrease by 1.33 PP if you are a high school graduate, 4.54 PP if you’ve taken a college class, and by 13.68 PP if you are a college graduate. Your odds of being a smoker increase by 6.24 PP if you’re unemployed, 3.89 PP if your income is between 10k and 20k, 2.46 PP if your income is between 20k and 35k, and decreases by 1.75 PP if you make over 35k. Compared to the baseline group your odds of being a smoker increase by 6.54 PP if you are in good health, 6.88 PP if you are in fair health, and 15.35 PP if you are in poor health. Our trend variable says for each additional year in the sample your odds of being a smoker decrease by .19 PP. It is important to note the coefficients on lncigaretteprice, ageisbetween25and34, ageisbetween45and54, ageisbetween55and64, highschoolgraduate, incomeover35k, isincomeunknown, and iyear are statistically insignificant.

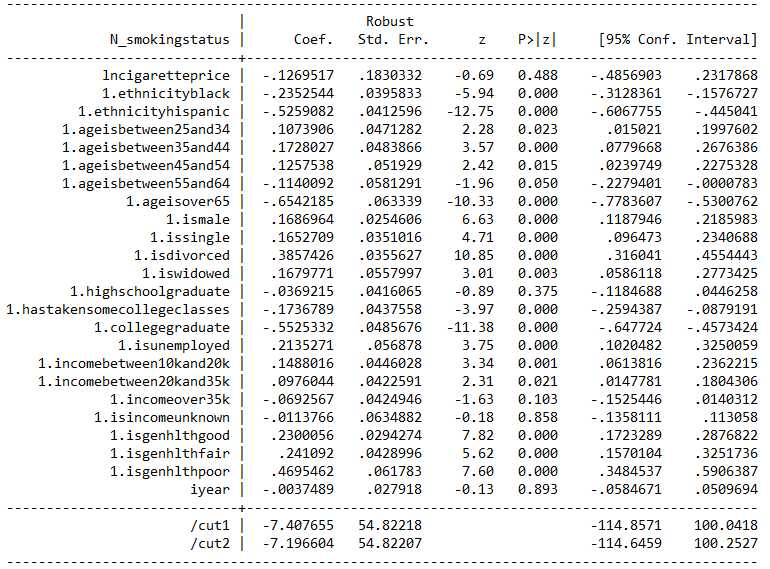


A marginal analysis of ZINB model of whether or not a person is a smoker reveals a lot about the participation effects of being a smoker. The ZINB model showed that among smokers the price elasticity was at the level of -0.4 and statistically significant. It indicates that for a 10% increase in price, the cigarette consumption per day would decrease approximately by 4.0%. Your predicted cigarette consumption is 1.22 cigarettes (cigs) lower if price goes up by 10%. Your predicted cigarette consumption decreases by 1.23 cigs if you are black, and 1.83 cigs if you are Hispanic. Your predicted cigarette consumption increases by 0.34 cigs if you are between 25 and 34, 1.05 cigs if you are between 35 and 44, 1.48 cigs if you are between 45 and 54, 1.33 cigs if you are between 55 and 64, and 1.06 cigs if you are over 65. Your predicted cigarette consumption increases by 0.52 cigs if you are male. Your predicted cigarette consumption decreases by 0.11 cigs if you’re single, 0.25 cigs if you are widowed, and increases by 0.33 cigs if you are divorced. Your predicted cigarette consumption if you’re a highschool graduate increases by 0.21 cigs, and decreases by 0.003 cigs and 0.61 cigs if you’ve taken some college classes, or are a college graduate, respectively. Your predicted cigarette consumption increases by 0.29 cigs if you’re unemployed. Your predicted cigarette consumption increases by 0.03 cigs if you earn between 10k and 20k annually, and decreases by 0.05 cigs if you earn between 20k and 35k, decreases by 0.15 cigs if you earn over 35k, and increases by 0.1 cigs if your income was unreported. Your predicted cigarette consumption increases by 0.37 cigs if your health is good, 0.44 cigs if your health is fair, 0.66 cigs if your health is poor, and 0.245 cigs for each year that has gone by. It is important to note the coefficients on ageisbetween25and34, issingle, highschoolgraduate, hastakensomecollegeclasses, isunemplyed, incomebtween10kand20k, incomebtween20kand35k, incomeover35k, isincomeunknown, isgenhlthpoor, and iyear are statistically insignificant.

It is important to note that the ZINB model uses a logit too, but the signs of our coefficients should be reversed given that we selected for issmoker compared to Mei’s selection of isnonsmoker. Mei also arrived at a price elasticity of -0.459, while we arrived at an elasticity of -0.4. Our findings that Blacks and Hispanics smoke less are in line with Mei’s. We can conclude that an increase in age from 18 and upward saw a statistically insignificant climb, with a sharp statistically significant decline in smoking prevalence at ages over 65. We both found a decrease in smoking prevalence from the base group of “Couple”, but Mei only had isdivorced as significant, while we only had issingle as insignificant. Mei’s paper had a similar trend of contradictory education participation effects, and weak effects on marginal cigarette consumption. Income effects between Mei and our findings were strongest, with extremely similar marginal effects on consumption and participation rates (small increase between income brackets, then sharp decline at over 35k, along with insignificant and minimal incomeunknown effects). We both find unemployed people are more likely to smoke, but smoke similar amounts of cigarettes to employed individuals. Our findings match Mei’s findings on health. People are more likely to smoke until it simply becomes a non-option, or the cost of doing so far outweighs the benefit (they have the poor health rating).

We were able to confirm Mei’s findings that TobaccoCPI is a poor predictor of smoking behavior. Smoking is highly addictive, and the choice to engage in smoking doesn’t seem contingent on price. Its effect on cigarette consumption had a statistically insignificant p-value, but still strongly presented itself in trend data on people who already smoked. The intensity effect seemed to present itself strongly. People seemed more likely to smoke when unemployed, or in compromised health points to the fact that smoking can be highly addictive. Mei’s paper could not conclusively say anything about relationship status and smoking, but our slightly larger sample had explanatory value. Our baseline group (Couples) was less likely to smoke, while emotionally comprised group like widows were. Smoking can be a strong emotional relief and relaxant, and being in a relationship seemed to be correlated to not smoking. It is possible having a partner provides a strong emotional release, similar to smoking. It may have helped to have CPI data specific to those counties. There are many omitted variables specific to California that could have caused large movements in error and heteroskadicity, which we failed to account for since we used national price level. This could have compounded with the fact that we added year as a trend variable instead of studying the fixed effects of being in California and adjusting the model to account for these factors.

Original idea

 We wanted to explore what entices one to change smoking status if the effect of price on smoking status is not significant. We constructed a new variable called N\_smokingstatus where 0, 1, and 2 were designations for nonsmokers, irregular smokers, and regular smokers, respectively. We ran an ordered probit on N\_smokingstatus and found that a 10% increase in price caused a 0.29 PP decrease in being a regular smoker, with a p-value of 0.49. We already established price is a statistically insignificant predictor of smoking status and cigarette consumption, but price also seem to have no statistically significant impact on whether people become new smokers, or more regular smokers. We then ran a multinomial logit and found a similar effect, stating price only decreasing your odds of smoking if you were already a heavy smoker. We theorized that not including price would help the fit of the ordered probit, but a likelihood ratio test between the null that the model fits better without price (restricted) vs a model with price(unrestricted) allowed us to reject the null with a p-value of 0.49 and chisquared of 0.48. This means that even though the margins of our ordered probit may not be statistically significant or accurate, our model predicts probabilities closer to the observed values with price than without. We’re running an ordered probit to see what makes people more likely to be nonsmokers because price only seems to be a deterrent to people who are already regular smokers.



Predict 1 = Non Smoker Predict 2 = Irregular Smoker Predict 3 = Regular Smoker

An analysis of the margins of our ordered probit reveal the same statistically insignificant 0.34 PP increase of being a nonsmoker for every 10% increase in cigarette price. Being black increases your chance of being a nonsmoker by 6.0 PP, while being Hispanic increased your chance of being a nonsmoker by 12.32 PP. Being between the ages of 25 and 34 decreased your chance of being a nonsmoker by 3.0 PP, while being between the ages of 35 and 44 decreases your chance of being a nonsmoker by 4.8 PP. Being between the ages of 45 and 54 decreases your chance of being a nonsmoker by 3.5 PP, while being between the ages of 55 and 64 increases your chance of being a nonsmoker by 3.0 PP. Being over the age of 65 increases your chance of being a nonsmoker by 14.77 PP. Being a male decreases your chance of being a nonsmoker by 4.63 PP. Being single, divorced, or widowed decreases the chance of being a nonsmoker by 4.64 PP, 11.52 PP, and 4.77 PP, respectively. Being a highschool graduate, college student (hastakensomecollegeclasses), or a college graduate increases your chance of being a nonsmoker by 1.0 PP, 4.58 PP, or 13.73 PP, respectively. Being unemployed, having an income between 10k and 20k, or having an income between 20k and 35k decreases your chance of being a nonsmoker by 6.20 PP, 4.19 PP, or 2.70 PP, respectively. Having an income of over 35k or an unreported income increase your chance of being a nonsmoker by 18.76 PP, or 3.08 PP, respectively. Having good health, fair health, or poor health decreases your chance of being a nonsmoker by 6.48 PP, 6.97 PP, or 14.55 PP, respectively. Our year trend variable states your chances of being a nonsmoker increase by 1.02 PP for each year that passes. It is important to note the effect of lncigaretteprice, highschoolgraduate, incomeover35k, incomeunknown, and year are statistically insignificant

Conclusion

Relative to Mei’s original paper, our findings do seem to follow similar trends regardless of the huge dispersion of our data for CPI compared to hers. We found that even though that a 10% increase in price in any model, probability or count, was not only insignificant but miniscule. Also, we found a very small effect of an increase in price of cigarettes on regular smokers that would deter them from consuming tobacco products. In all models, family statuses such as whether a couple was together or divorced and if someone was widowed had statistically significant effects. This correlates with our belief that if you have a partner, they will provide an emotional relief similar to that of, or similar to smoking. Another interesting result that we found was that people with higher forms of education are less likely to smoke compared to their less educated counterparts. From our findings we can conclude that the largest deterrent of being a smoker is being over the age of 65. Finally, as years were increasing the results showed someone would be less likely to become a smoker.

Discussion

From the results we found based on smoking and education, this can also be related to income because it’s more likely for people with higher education to make more money and therefore will be less likely to be smokers. Some reasons we thought of as to why so many people were less likely to become smokers after the age of 65 was maybe due doctors monitoring their health and telling them they need to take care of themselves. Another reason would possibly be that after the age of 65 people who have not smoked their entire lives are probably not going to change up their habits that late in life. This also follows with the effects of the progression of age/year because as people get older their less likely to change their habits unless possibly some event causes psychological distress in class.

Further Research

The general trend is that the price level of tobacco products have been increasing for decades and smoking prevalence has gone done drastically. So, given more funding and the more recent BRFS surveys do these trends still hold given the explosion in the new market for tobacco products made primarily of vaping products? It would be worthy of future research to see if emotional well-being is a better predictor of smoking status than price.

References

Sung HY, Hu TW, Keeler TE. Cigarette taxation and demand: an empirical model. Contemp Econ Policy 1994; 12(3): 91–100.

Keeler TE, Hu TW, Barnett PG, et al. Taxation, regulation, and addiction: a demand function for cigarettes based on time-series evidence. J Health Econ 1993; 12(1): 1–18.

Hu TW, Sung HY, Keeler TE. Reducing cigarette consumption in California: tobacco taxes vs an antismoking media campaign. Am J Public Health 1995; 85(9): 1218–1222.

Hu TW, Ong M, Keeler TE, et al. California cigarette consumption: the effects of the Proposition 10 tax and national tobacco industry settlement. Report submitted to Tobacco Related Disease Research Program, University of California, 2001.

Hu TW, Ren QF, Keeler TE, et al. The demand for cigarettes in California and behavioral risk factors. Health Econ 1995; 4(1): 7–14.

Vuong Q. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 1989; 57: 307–334.