VILNIUS UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

LUKAS DRAZDYS

Quantitative Economics, 3rd year

Bachelor Thesis

EFFECTIVENESS OF TOBACCO TAXATION ON SMOKING PREVALENCE AND CONSUMPTION: A CASE OF ITALY

Student ________________________________

(signature)

Supervisor ________________________________

(scientific degree of the supervisor, academic title, name, surname)

____

(signature)

Date of submission of the academic paper

____

Registration No. __________________________

Assessment of the academic paper

____

(Date, assessment score, signature of the chairman of the commission)

Vilnius, 2021

Contents

Introduction 3
1. Literature 4
2. Methodology 7
2.1 Similarity matrix 7
2.2 Smoking prevalence 7
2.3 Cigarette demand 8
2.4 Model Selection Algorithm 9
2.5 Data description 11
3. Data analysis and Results 15
3.1 Similarity matrix 15
3.2 Smoking prevalence model 16
3.3 Cigarette demand model 21
4. Possible scenarios 26
4.1 Scenario 1 26
4.2 Scenario 2 29
4.3 Scenario conclusions 31
Conclusion 32
References 33
Summary 34
Annex 36

Introduction

Nearly 700,000 people die from smoking tobacco every year in Europe. According to Tobacco Atlas[1], from those 700,000 - 93,300 are killed in Italy by tobacco-caused diseases. Italy has introduced many policies to combat this tobacco epidemic. Italy provides national quitline services and helps people through nicotine replacement therapy and cessation services, for which the government covers costs. They also have plain packaging, with 65% of the pack covered in graphic images and/or warnings. Anti-tobacco mass media campaigns were also run to warn people about the dangers of tobacco.

Furthermore, multiple bans on tobacco advertising, both direct and indirect, were passed. In Italy, taxes for smoking are relatively high as well. Excise tax in Italy amounts to 58.01% of the final price, whereas the WHO benchmark is at a minimum of 70% of the total price. However, tobacco-related complications and deaths are still high. An increase in tobacco prices has been proven repeatedly to be an effective way of controlling the smoking population.

Multiple types of research are being conducted, how cigarette demand is affected at different income levels or how it is affected at different price levels. In more recent analyses[2], World Health Organization MPOWER indices were also included. They provide a lot of insight into each sector of tobacco control. Other studies provide country groups, where each group has its price elasticities. The pretext is that impoverished countries have higher elasticities and more developed countries lower elasticities.

Previous studies use various techniques to estimate price elasticities of cigarette demand: panel threshold models[2], 2SLS, and the Engle-Granger two-step method[3] and cross-sectional analyses[4]. However, not all studies accounted for omitted variable bias when using cigarette price as a regressor in their models. Price is entirely dependent on taxes, and by not including them in the model through 2SLS or some other way, the results of those studies are not completely accurate. One more drawback of other studies is that they mostly build only cigarette demand models, which only provide how much people smoke, not the smoking population. That might lead to misinterpretation of their final results. If most people are hard daily smokers, they would only reduce the amount of tobacco they consume and not quit altogether.

My contribution to tobacco control will be the creation of a government revenue tool that will show how a change in taxes responds to cigarette demand, smoking population, and government revenue. First, I will create fixed-effect panel models, using country-level data to find cigarette price elasticities, using a unique model selection algorithm. Secondly, I will focus on Italy and check how each driver is contributing to the model. To do so, driver growth decomposition will be constructed. Finally, I will create the government revenue tool specifically for Italy. This way, it will be a lot easier to understand what is happening when prices increase. How cigarette demand adjusts to the new fees and how well such a policy would affect smoking prevalence. And because Italy does not have much room to control smoking through bans, this tool can help make better decisions.

1. Literature review

I will be comparing a couple of research papers that have been conducted in the European region. In an article by Yeh, CY., Schafferer, C., Lee, JM. et al.[2], 28 EU countries were analyzed during the years 2005-2014. This article is an excellent benchmark for my models since the same data sources are used. A standard constant elasticity log-linear demand model was created. The dependent variable was annual cigarette consumption per capita, and explanatory variables were price per cigarette, national income, rural population, the highest amount of MPOWER indices implemented, and cigarette prices in Eastern European countries. Here an assumption was made that countries with different incomes would react to the model differently. Therefore, a panel threshold model was used with three thresholds. Those thresholds were chosen by per capita national income, giving that countries with the lowest federal income produce price elasticities higher than -1 and at the highest threshold of -0.503. Based on this article, Italy is part of the countries where price elasticity is -0.503. However, these results may not be that accurate since there are no instrumental variables, and the model is possibly suffering from omitted variable bias.

When looking at another research conducted by Lien Nguyen, Gunnar Rosenqvist, Markku Pekurinen[3] in 2012, they analyzed 11 countries using dynamic models with 2SLS estimation Engle-Granger two-step method, they have found that Italy has lower elasticity. Here the main explanatory variables were tobacco prices, past and future values of the dependent variable (cigarette demand), and some other variables that could explain the change in demand. Having dynamic models helps the author capture some addiction factors in his study, since immediate price increases might not instantly affect consumers. Time-series data were used, therefore, additional autocorrelation tests were performed to justify the validity of the models. The lowest short-run elasticity is only -0.246, and the highest long-run elasticity is -0.418, which could be interpreted as 2SLS reducing the price elasticity. In other countries, the elasticities would be reduced even more, sometimes even below -1, which means that an increase in cigarette price would reduce the demand by more than the government would gain revenues, i.e., revenues would reduce even though prices increase.

One more study was conducted using older data, from 1970 to 2001, called the Economic aspects of smoking in Europe[4]. Three different logarithmic regressions were built to see how much of the demand coefficients are explained by other variables. However, very few variables were tested. Only price, female smoking prevalence, and the degree of legislation are dummy variables.

The elasticities ranged from -0.47 to -0.67. Looking specifically at the case of Italy, national time-series analyses were conducted for smoking prevalence. A logarithmic regression was also used, and only two independent variables were used, price and GDP per capita. Here the price elasticity for prevalence was -0.3 between 1970 – 2001. However, since many years were missing in the data, it was more likely that the elasticity is higher, about -0.37. This is the only paper discussed that took the smoking population into account. Overall, looking at the various researches, data samples, and methods, it is clear that cigarette price is an effective way to increase tax revenue and reduce smoking. The increased revenues could also contribute to further investment in reducing smoking prevalence. In Table 1 a literature summary is presented, with the key findings that will be later compared with my findings and the overall scope of each research.

Table 1: Literature Summary

Author	Sample	Scope	Methodology	Findings for Italy
Chun-Yuan Yeh, Christian Schafferer, Jie-Min Lee	2005-2014	28 EU countries	Panel Threshold Regression	The price elasticity of demand is -0.503
Lien Nguyen, Gunnar Rosenqvist, Markku Pekurinen	1953-2009	11 EU countries	Dynamic models with 2SLS estimation	Price elasticity of demand ranges from -0.246 to -0.439
Dr. Esteve Fernández	1970-2001	22 EU countries	Cross-sectional analysis and logarithmic regressions	Price elasticity of demand is -0.43, and elasticity for prevalence is -0.30
David Levy, Silvano Gallus, Kenneth Blackman[5]	1999-2010	Italy	SimSmoke model: population model, a smoking model, a SAD model, and policy modules	Elasticities for smoking prevalence from previous researches were reused for different age groups: 15-24 elasticity -0.3, 25-34 elasticity -0.2, 35+ elasticity -0.1
Elaine Fuertes, Alessandro Marcon, Laura Potts[6]	2018-2058	Italy	DYNAMO-HIA model	Elasticity for smoking prevalence is -0.5

2. Methodology

In this section, I will discuss the methods used to generate and create the models used for further analysis. The final models that will be used for further analysis will be developed using a specially made algorithm. Therefore, all potential drivers will be discussed. First of all, since there is a small number of years to estimate the effects of the drivers correctly, multiple countries need to be used in the estimation process. To do so, panel models will be used. Then the question remains, what countries to use for the estimation since even looking at the European countries, they are all different in population, living standards, etc., a similarity matrix is constructed to see which countries are similar by specific categories. After selecting the countries, I can start building the models which will not only show what drives consumption but also will give the price elasticities. They are crucial for my analysis to help me make the government revenue tool because prices will be the primary variable that will be changed.

2.1 Similarity Matrix

A similarity matrix is a table that shows how each element in it is similar to other aspects of the table. In this case, European countries will be used to see which are the most identical by key indicators. Since using all the data or only one category is too unspecific, a particular set of key indicators is used. Multiple categories are chosen that are specific to this research and then compared for each country, constructing the similarity matrix. I used a similarity index for comparing coupled matrices method by Ulf G. Indahl, Tormond Næs, and Kristian Hovde Liland[7]. According to the authors, the SMI (Similarity of matrices index) method is better. It deals better with larger matrixes and other coefficients such as RV[8] and RV2[9], introducing bias with increasing samples. The overall method compares a subspace of combinations derived by matrix decomposition strategies like principal component analysis and partial least squares regression. This coefficient is being used, because the primary tool for the analysis (RStudio) has a package implemented by the papers' authors and works excellent for need. Using this coefficient, I get exact percentages of how similar each country is.

2.2 Smoking Prevalence

To calculate the potential effects on the smoking population, I use fixed-effects panel models. The model selection algorithm is run based on the following notation:

$$SmokingPrevalence_{it} = \alpha_i + \beta_1 price_{it} + \beta_2 X_{2it} + \cdots + \beta_6 X_{6it} + u_{it} \quad (1)$$

Where *SmokingPrevalence* is the dependent variable, α is the country-specific fixed-effect, $price$ is the logged average price of a cigarette stick and X_2 through X_6 are additional drivers that may help explaining smoking prevalence. β_1 through β_6 are the slopes or elasticities of each variable, and u is the country-specific zero-mean random-error term.

The advantage of using panel models instead of pure single country time series models is due to higher variability and efficiency. Additionally, panel models can provide significantly better results when using a smaller sample size. Implementing a more heterogenous approach also helps to reduce omitted variable bias, which could arise from cigarette price. At the same time, this could potentially be a drawback, since we do not know which biases the model is omitting. It could be cigarette taxes that directly affect their price, or something completely different. Nevertheless, it is still better to use fixed-effect panel models, because we do not need to define all omitted variables compared to pooled OLS. These types of models also provide larger standard error compared to pooled OLS, which could also play a part in having lower statistical power. However, in this study, the pros strongly outweigh the cons and it is still better to use fixed-effect panel models.

2.3 Cigarette demand

The cigarette demand model is similar to the one for smoking prevalence. However, here a two-stage regression is needed. Our price variable has a lot of unobserved effects on cigarette demand. Therefore, the β of this coefficient needs to be adjusted with an instrumental variable. There are two possibilities for an instrument, a tax that affects cigarette price or own price lag. A better option here would be to use the lag of cigarette price to estimate the correct beta of cigarette price in the final fixed effects panel model regression, because some taxes in some countries are different for each cigarette quality. Since the prices I use are the average price per stick, there is no way to account for that. That leaves only VAT tax, which is the same for all cigarettes, no matter their price. Unfortunately, the VAT in some countries does not change or is inversely correlated with the price variable. Therefore, cigarette own price lag is the best available option to estimate the final regressions' beta accurately. The beta is calculated with the following formula:

$$\hat{\beta}_{IV} = (Z'X)^{-1}Z'Y \quad (2)$$

Z is the matrix of lagged cigarette price, X is the matrix of cigarette price, and Y is the matrix for cigarette demand. With the beta estimated, a proper panel model regression can be constructed.

Again, it will look the same as the one in the smoking prevalence model (equation 1). Only the beta for cigarette price will be the new one.

$$SmokingDemand_{it} = \alpha_i + \hat{\beta}_{IV} price_{it} + \beta_2 X_{2it} + \cdots + \beta_6 X_{6it} + u_{it} \quad (3)$$

Now that the beta for cigarette price is adjusted, the proper regression can be constructed.

2.4 Model selection algorithm

The process of selecting the correct models is done by a written algorithm, which combines all the potential drivers and tries fitting them into the model. There are 58 drivers, and the cigarette price has to be in each of the models. A total of six or seven drivers will be used in the models. For example, if the price is one of the drivers and there will be five additional drivers, that gives a total of 458,377,920 (56 * 55 * 54 * 53 * 52 * 1) model combinations. The sample years are from 2007 to 2016. Then years from 2017 to 2019 are treated as out of sample years. And from 2020 to 2024 are forecast years, built by Euromonitor analysts. They are used as a benchmark to see how my forecasts compare to Passport forecasts. These forecast years are later used in the government revenue tool. First, the algorithm takes the sample data and tries to fit the model and then checks if the signs are as expected in the driver list. Additionally, the p-value is also checked. This is repeated for all the possible combinations, and all the potentially good models are outputted in a separate file. Here I can see the overall model, its MASE (mean absolute scaled error), MAPE (mean fundamental percentage error), and country-specific MASE and MAPE statistics to see how well the models fit. Moreover, the same statistics for the Out sample years are also calculated. Alongside all this, an additional file is created with all the model graphs to see visually how well the model looks because the statistics aren't enough, and having a solid statistic does not mean a good overall fit. All the models are then mechanically checked to see what drivers work or ruin the models. Then the driver list is cleaned to remove problematic or unused drivers, speed up the whole process, and look for a better/best model. Looking at the drivers of the fitted models, I can also check if there are no "duplications", two or more different drivers that explain the same meaning in the model. For example, if a model consists of minimum wage per hour and disposable income, this would be bad because both these drivers give the same meaning. Higher minimum wages mean more disposable income, which means more money for non-essential purchases like cigarettes and higher smoking prevalence. Both are still included and checked, because one might give better explanatory value to the final model.

Table 2: Data Preparation and Algorithm Summary

	Subsection	Process
Data Preparation	Data cleaning	All Passport data is being cleaned for all 80 countries and prepared for modeling use. The transformation from wide to long format, preliminary category selection is run.
	Additional data cleaning	Additional data from WHO is cleaned and added to clean Passport data.
	Data standardization	All data is transformed to constant prices and 2019 fixed exchange rates for more straightforward interpretation. Variables are converted to per capita terms and logged.
	Forecasting	Categories that do not have data for 2020-2024 naïve forecasts are produced to avoid relevant models not being calculated.
	Setup file creation	Additional files are being created and prepared for modeling. Final data file expected signs file.
Modeling	Similarity Matrix	A similarity matrix is produced, and top countries are selected to be included in the model based on the land of interest.
	Setup	All data is read, inputs for in-sample and out-of-sample years are created, p-value threshold and the number of drivers to be used are also specified. Finally, selecting what will be modeled and countries from the similarity matrix are introduced.
	Data Filtering	All data is then filtered based on the expected signs file, and countries form a similarity matrix.
	Modeling Algorithm	Combinations of drivers in the expected signs file are created based on the number of drivers in a single combination. Each combination is then tested, and the model's output is checked with the p-value threshold and expected signs. If everything is as expected, the model summary with MASE and MAPE statistics is placed in a separate file to be checked later, alongside model graphs (actual values versus fitted).
	Result checking	Model results are checked mechanically to see if drivers are logical, and their coefficients are not too high. If everything is good, the best model is selected, and then I move on to the finalization. If there are no good models, the number of drivers used to generate combinations is reduced, and the p-value threshold is adjusted to include more models. The exact process is repeated until a good model is produced.
Finalization	Graph creation	With the best models selected, final graphs are created: model actual versus fitted graph and growth decomposition.

I would also like to discuss why these accuracy measure statistics were chosen. First, MASE gives an excellent threshold to check if the model is built accurately. A MASE of 1 says that it would be the same as using the change in previous years to predict the difference in the future years. Or in other words, the forecast follows the naïve random process of in-sample data. Therefore, looking at models with the MASE of less than 1 gives a better prediction than simply using naïve forecasting, and this tells me that the model correctly predicts future values. Second, MAPE is another accuracy measure to see how off the out-of-sample forecast of the model is. It gives more flexibility in choosing a well-fitted model since it only shows the percentage of how off the forecast is from actual values. Having these two statistics provides a good understanding of how well the models are being fitted.

2.5 Data description

Data used is from two sources. The main is from Passport (Euromonitor International database) and World Health Organization database. The only reason the WHO database is included is that Passports smoking prevalence data is not that great. For some countries, the trends are correct, however when using a sample of 35 European countries, the differences matter. Moreover, WHO has also had MPOWER[10] indices, representing different ways to control tobacco consumption (See Annex 1). MPOWER is an acronym, which translates to each of the indexes: Monitor, Protect from tobacco smoke, Offer help to quit tobacco use, Warn about dangers of tobacco, Enforce bans on tobacco advertising and Raise taxes on tobacco. However, there is one issue with WHO data, MPOWER indices are only for the 2007-2018 timeframe, and only every other year has actual data points. Interpolation method where the same value as a historic year is repeated until there is a change in the index until the final forecast year. WHO smoking prevalence also has some missing years, so a similar interpolation method was introduced. The lost years had the value between the known years and followed the same trend, and the whole series and forecast years (including 2019) were added using the naïve method with a drift. Passport data that is used is confidential, and due to this, I cannot provide a detailed description of the data. It consists of different category types: alcoholic drinks, economy, employment, expenditure (consumer and government), households, income, living standards, population, and tobacco. Having a wide range of categories and industries provides better explanatory value for the models. The different categories alongside the expected signs can be seen in Table 3. The expected signs are pretty self-explanatory, 1 means a positive slope/elasticity is expected, -1 a negative one, and 0 that it can be either. Definitions for the variables included in the modeling process can also be seen in Annex 1.

The data used is annual, from 2005 to 2024. Years 2005 through 2019 are historical years, and 2020 - 2024 are forecasts calculated and built by Euromonitor analysts. The historical years are in current prices and year-on-year exchange rates, and forecast years are in 2019 constant prices and 2019 fixed exchange rates, which is later transformed to where all data is in 2019 constant prices and 2019 fixed exchange rates. Some categories have missing forecasts, so using the naïve method, forecast years are produced for each category: minimum wage per hour, wage per hour, human development index. For modeling, the data is not being used as-is. Instead, all the variables that are not already in percentages or shares are transformed to per-capita, logged or made per capita, and then logged. At first, all drivers are being used to estimate the models. However, the list is later reduced after a few iterations of modeling. A list of all the drivers and expected signs can be seen in Table 3.

Table 3: Driver List and Expected Signs

Industry	Category	Sign	Source
Alcohol	Alcoholic Drinks Price	-1	Passport
Economy	Real GDP Growth	1	Passport
Economy	GDP	1	Passport
Employment	Employment Rate	1	Passport
Employment	Male Employment Rate	1	Passport
Employment	Female Employment Rate	1	Passport
Employment	Unemployment Rate	0	Passport
Employment	Male Unemployment Rate	0	Passport
Employment	Female Unemployment Rate	0	Passport
Expenditure	Consumer Expenditure on Fruit	-1	Passport
Expenditure	Consumer Expenditure on Alcoholic Beverages and Tobacco	1	Passport
Expenditure	Consumer Expenditure on Alcoholic Drinks	1	Passport
Expenditure	Consumer Expenditure on Water and Miscellaneous Domestic Services	-1	Passport
Expenditure	Consumer Expenditure on Electricity Gas and Other Fuels	-1	Passport
Expenditure	Consumer Expenditure on Education	-1	Passport
Expenditure	Government Expenditure on Education	-1	Passport
Expenditure	Government Expenditure on Health	-1	Passport

Table 3: Driver List and Expected Signs (continued)

Industry	Category	Sign	Source
Household	Average Household Size	0	Passport
Household	Average Number of Children Per Household	1	Passport
Household	Average Size of Urban Household	0	Passport
Household	Average Size of Rural Household	0	Passport
Household	Number of Households	0	Passport
Household	Households with 1 Child	0	Passport
Household	Households with 2 Children	0	Passport
Household	Households with 3 Children	0	Passport
Household	Households with 4 and more Children	0	Passport
Income	Minimum Wage per Hour	1	Passport
Income	Wage per Hour in Manufacturing	1	Passport
Income	Wage per Hour	1	Passport
Income	Disposable Income	1	Passport
Life Standards	Percent of Population Aged 15 Plus with Higher Education	-1	Passport
MPOWER	Monitor	-1	World Health Organization
MPOWER	Protect from Tobacco Smoke	-1	World Health Organization
MPOWER	Offer Help to Quit Tobacco Use	-1	World Health Organization
MPOWER	Warn About the Dangers of Tobacco	-1	World Health Organization
MPOWER	Enforce Bans on Tobacco Advertising	-1	World Health Organization
Population	Dependency Ratio	1	Passport
Population	Death Rates	-1	Passport
Population	Deaths	-1	Passport
Population	Life Expectancy at Birth	0	Passport
Population	Net Migration Rate	0	Passport
Population	Generation Z	-1	Passport
Population	Millennials	0	Passport
Population	Generation X	0	Passport

Table 3: Driver List and Expected Signs (continued)

Industry	Category	Sign	Source
Population	Teens Aged 13-17	0	Passport
Population	Young Adults Aged 18-29	0	Passport
Population	Middle Youth Aged 30-44	0	Passport
Population	Mid-lifers Aged 45-64	0	Passport
Population	Later Lifers Aged 65-79	0	Passport
Population	Seniors Aged 80 Plus	0	Passport
Population	Urban Population	-1	Passport
Tobacco	Cigarette Price	-1	Passport
Tobacco	Cigarillos Price	1	Passport
Tobacco	Fine Cut Tobacco Price	1	Passport
Tobacco	Smoking Tobacco Price	1	Passport
Tobacco	Cigars Price	1	Passport
Tobacco	Pipe Tobacco Price	1	Passport
Economy	Human Development Index	-1	Passport

3. Data analysis and results

This section will present all the results, and a deeper look into the described methodology will be given. The similarity matrix, model selection process and model results, and more profound analysis of each driver used in the selected model.

3.1 Similarity matrix

The similarity matrix consists of 35 European countries that exist in the data (Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Netherlands, North Macedonia, Norway, Poland, Portugal, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom). The critical indicators chosen to calculate the similarity percentages are the demography, economy, living standards, and tobacco industries, which are pretty wide to select the countries for the regression correctly. From the demographic part, the urban population percentage of total population and unemployment rate is chosen. These indicators capture the level of urbanization and the rate at which unemployment is increasing/decreasing in the selected countries. The economy part consists only of GDP per capita, which represents the economic status quite well. The more important criteria are those of life standards. Percent of the population with higher education and consumer expenditure on electricity, gas, and other fuels, share in consumer expenditure, represents living standards in each country. The most important aspect of my research is the tobacco industry, which is defined by cigarette retail volume per capita in the similarity matrix. The top 18 countries (that are the most similar to Italy) are selected.

Table 4: Similarity Matrix Results for Italy

	Italy		Italy
Austria	60.8%	Lithuania	59.4%
Belarus	57.9%	Netherlands	67.4%
Belgium	71.9%	North Macedonia	60.0%
Bosnia and Herzegovina	65.6%	Norway	63.0%
Bulgaria	70.2%	Poland	68.5%
Croatia	61.3%	Portugal	60.7%

Table 4: Similarity Matrix Results for Italy (continued)

	Italy		Italy
Czech Republic	62.7%	Romania	58.6%
Denmark	53.3%	Russia	61.6%
Estonia	78.0%	Serbia	67.7%
Finland	61.9%	Slovakia	66.3%
France	66.7%	Slovenia	62.1%
Georgia	65.5%	Spain	70.3%
Germany	67.9%	Sweden	78.5%
Greece	66.6%	Switzerland	54.6%
Hungary	63.1%	Turkey	60.3%
Ireland	65.4%	Ukraine	45.2%
Italy	100%	United Kingdom	68.3%
Latvia	64.0%		

For this paper, I will be looking into Italy, so as seen in the similarity matrix, the list of countries to be used in our panel model regression are Italy, Sweden, Estonia, Belgium, Spain, Bulgaria, Poland, United Kingdom, Germany, Serbia, Netherlands, France, Greece, Slovakia, Bosnia and Herzegovina, Georgia, Ireland, and Latvia, in descending order of similarity, from Sweden being 78.5% and Latvia being 64% similar to Italy.

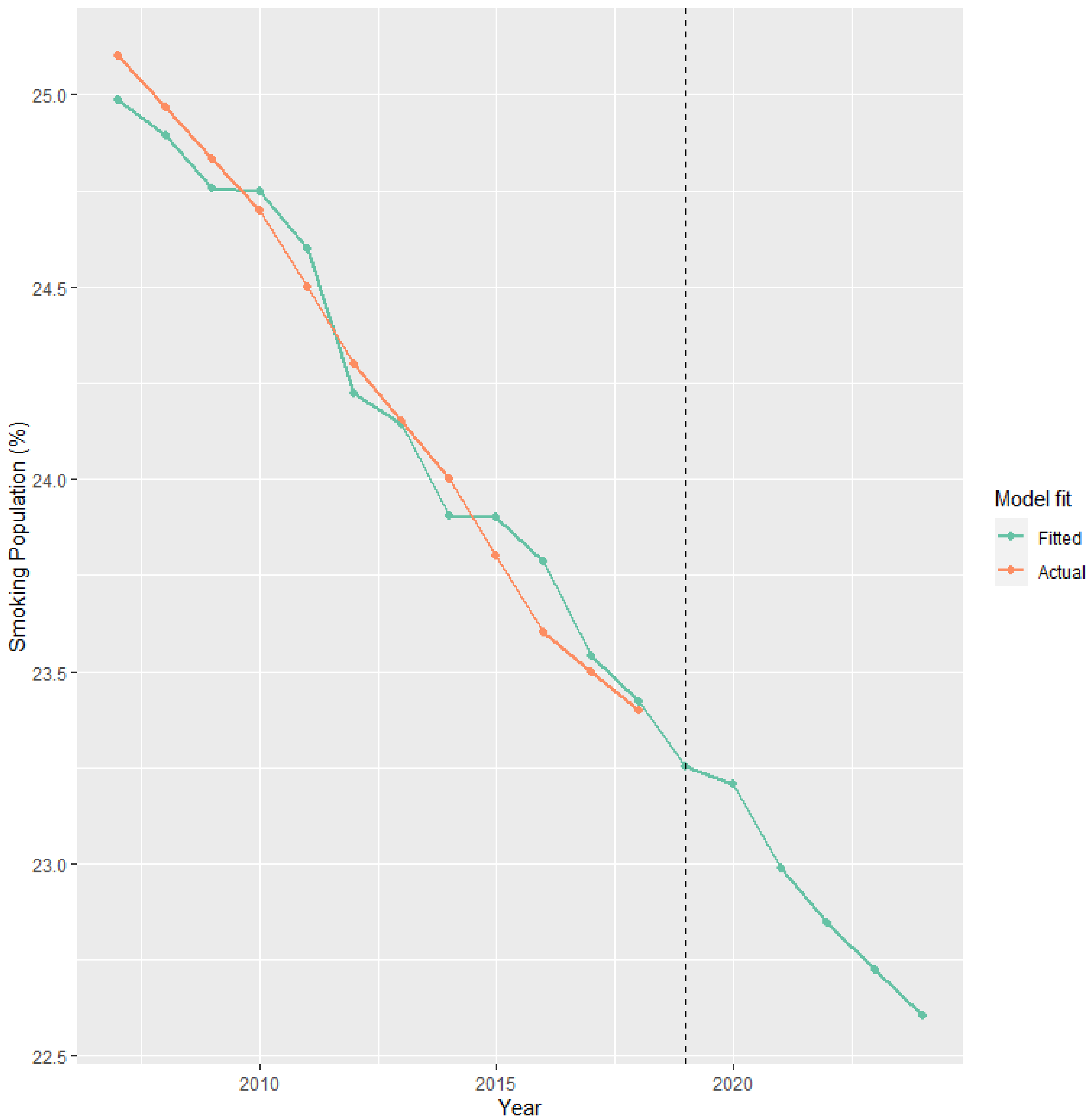
3.2 Smoking prevalence model

To find the best model of smoking prevalence, it took few iterations of the modeling process. The model chosen can be seen in Table 5.

Table 5: Smoking Prevalence Model

driver	Coefficient	Std. Error	t-value	Pr(>|t|)
log.wage.per.hour	0.2850	0.0589	4.843	0.0000
log.consumer.expenditure.on.water.and.miscellaneous.domestic.services.cap	-0.0553	0.0188	-2.941	0.0038
log.government.expenditure.on.health.cap	-0.0826	0.0318	-2.600	0.0104
percent.of.population.aged.15.plus.with.higher.education	-0.4657	0.2515	-1.851	0.0663
human.development.index	-0.8383	0.3616	-2.318	0.0219
log.cigarette.price	-0.1703	0.0364	-4.676	0.0000

Figure 1: Smoking Prevalence in Italy

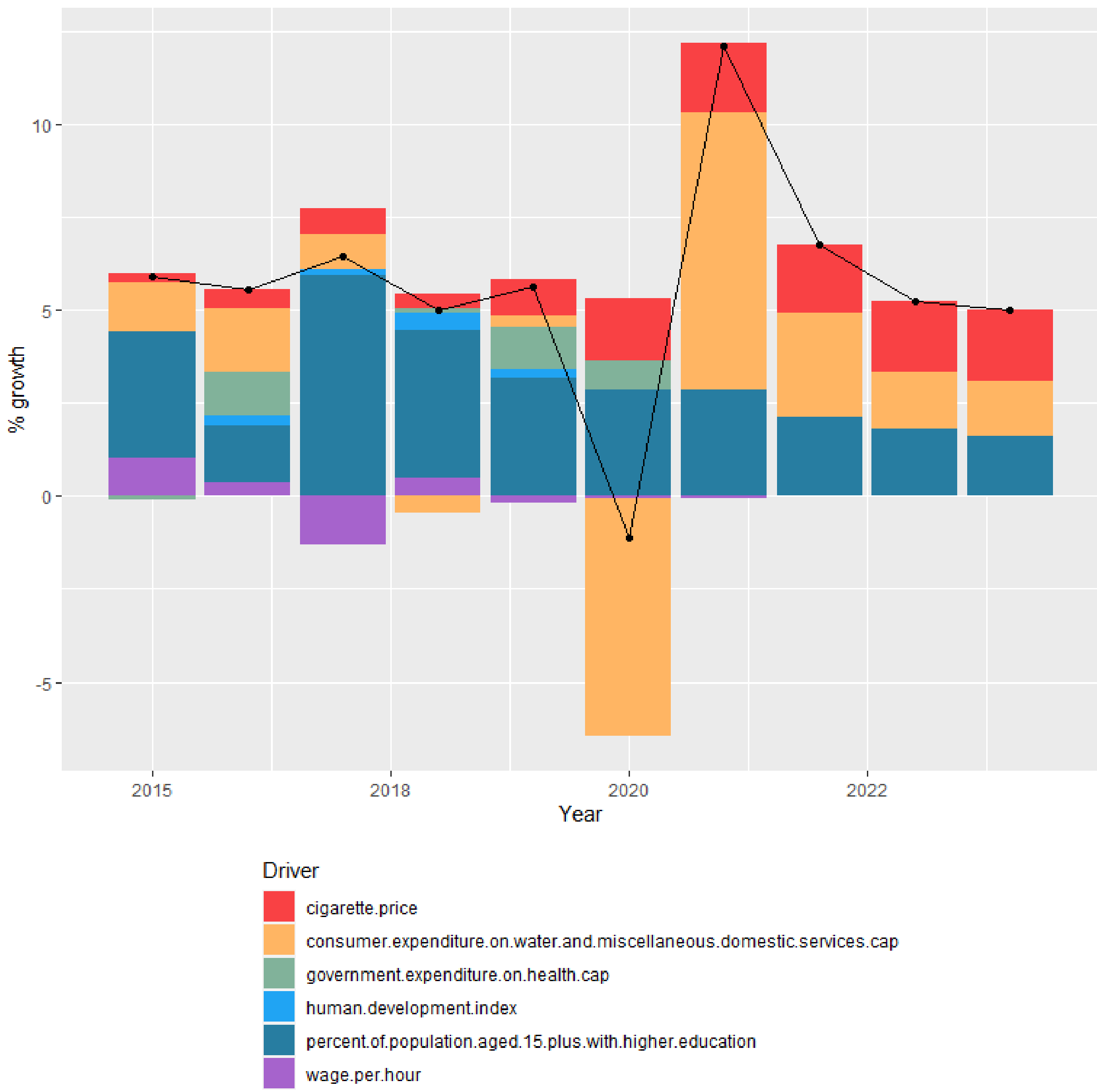


Graphically, the fit could be better. However, the main interest is the forecast years. The out of sample MAPE for Italy is only 0.0820% which is very good. A good benchmark for MAPE is 15%, and now it is a lot better. The MASE in the out of sample years for Italy is 0.2470, which is also relatively lower than the threshold value. In-sample MAPE is 0.2558%, and MASE is 0.5278. Overall results are pretty satisfying and with logical drivers. See Annex 1 for definitions.

- Log wage per hour and log consumer expenditure on water and miscellaneous domestic services say that increasing wages gives more disposable income, which allows people to smoke more, and expenditure on miscellaneous domestic services reduces that disposable income and therefore reduces prevalence (more in growth decomposition).

- Log government expenditure on health per capita tells me that when expenditure increases, prevalence also decreases due to a probable increase in help for smokers to quit.
- The percentage of the population with higher education also reflects negatively on smoking prevalence. More educated people are more aware of the dangers of tobacco and are less likely to smoke.
- The human development index coefficient is relatively high. However, a negative sign is to be expected because of how that index is constructed. I expect to see a negative sign because the higher the index, the more developed the people there are, which means longer life span, healthier population (not necessarily meaning, non-nonsmokers), from the aspect of "affording" to ruin their health by things like smoking.
- And finally, cigarette price shows that price negatively affects the smoking population and negates the growth of new smokers. The coefficient can be interpreted as the price elasticity of demand. In this case, a price increase of 10% will result in a 2.12% decrease in smoking prevalence.

Figure 2: Smoking Prevalence growth decomposition for Italy



Looking at the growth decomposition, what stands out is consumer expenditure on water and other miscellaneous domestic services. Since it has a negative sign in the model, expenditure increased in 2020, probably due to covid-19 related reasons. It reflects negatively on smoking prevalence because of people having less disposable income. And then expenditure decreases in 2021. Such changes in this variable are most likely due to home seclusion during covid-19. The percentage of the population with higher education also takes up quite a lot in the growth decomposition. It reflects that every year there were less and less people who have higher education, thus increasing prevalence overall. One more thing visible from the growth decomposition is that cigarette price was constantly growing, following smoking prevalence.

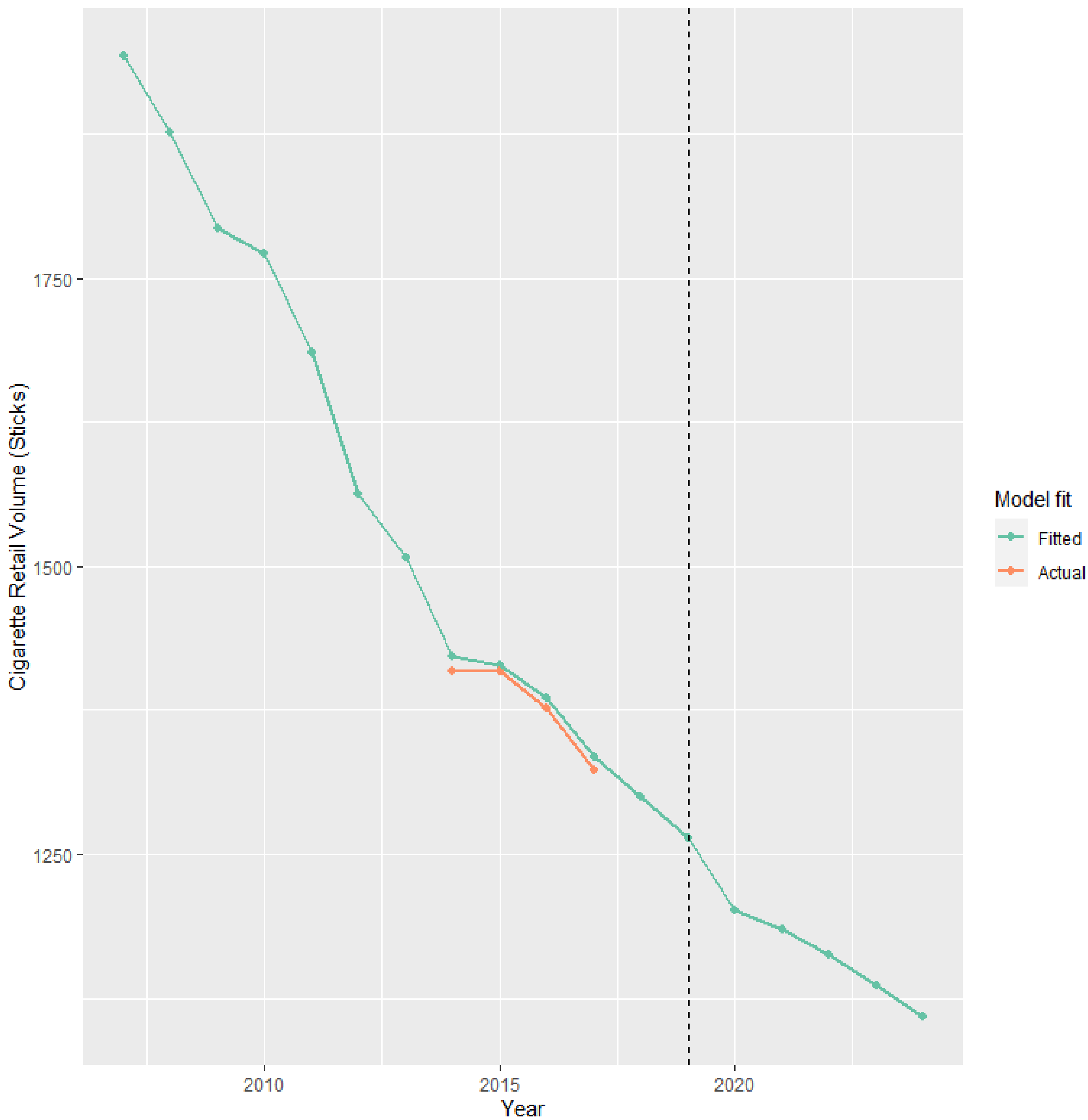
3.3 Cogarette demand model

When looking for the smoking demand model, it was pretty hard to find a model with price elasticity in line with other studies.

Table 6: Cigarette Demand Model

driver	Coefficient	Std. Error	t-value	Pr(>|t|)
average.household.size	0.5521	0.2023	2.730	0.0071
human.development.index	-2.2241	0.9334	-2.383	0.0184
urban.population.cap	-1.8188	1.1500	-1.582	0.1158
log.consumer.expenditure.on.alcoholic.drinks.cap	0.2215	0.0599	3.695	0.0003
log.wage.per.hour	0.7382	0.1099	6.718	0.0000
log.cigarette.price	-0.8362	0.0860	-9.719	0.0000
enforce.bans.on.tobacco.advertising	-0.1304	0.0653	-1.998	0.0475

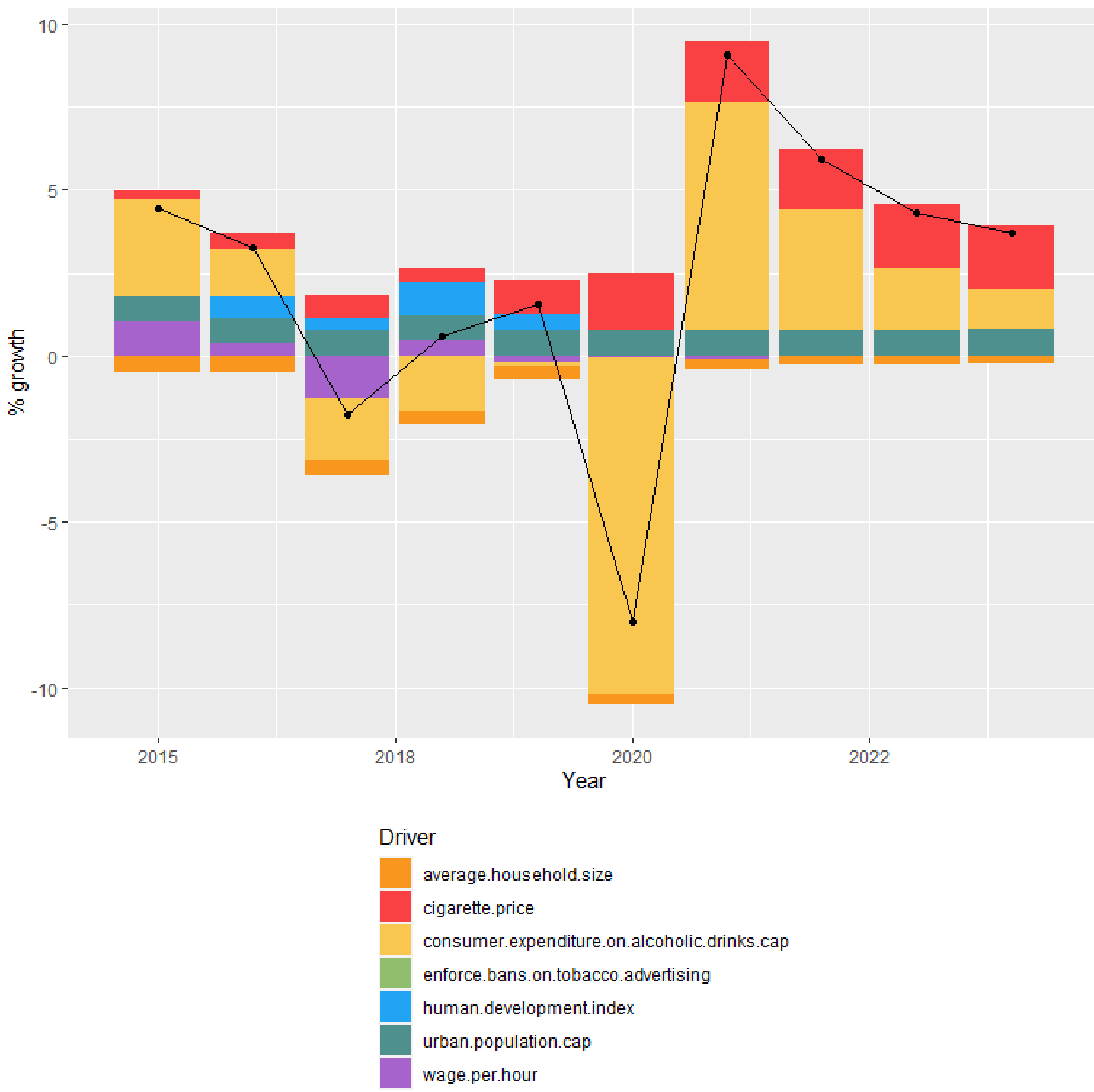
Figure 3: Cigarette Retail Volume in Italy



The fit of the model is very well. However, I can only provide a small chunk of actual cigarette retail volume data due to data confidentiality. Nevertheless, the out of sample MAPE for Italy is 0.1271%, and MASE is 0.3289, which tells me that projections are robust. The in-sample statistics for Italy are MAPE at 0.1500% and MASE at 0.3183, which are pretty close to the statistics in out of sample years. And looking at the whole picture, the fit is almost perfect. The drivers could probably tell a better story about why this was such a great fit.

- Average household size has a positive sign in the model. The interpretation here is that with more people in a household, the likelihood of someone being a smoker is higher than for a smaller household. Thus, if one of the adults is smoking, the other probably also smoke.
- Here, the human development index has a higher coefficient than in the prevalence model, hover the interpretation does not change.
- The percentage of the population living in urban areas has a negative sign in my model because people in urban areas are less likely to be smokers than in rural areas.
- Alcoholic drinks complement cigarettes; thus, increased expenditure on alcoholic beverages also increases cigarette demand.
- Wage per hour is the same as with the prevalence model: disposable income increases, more money for cigarettes.
- Enforce bans on tobacco advertising is an MPOWER index, representing how many media channels or forms of tobacco advertisements are banned. If tobacco is not advertised as much, fewer people feel the need to smoke.
- And finally, cigarette price has a more significant negative contribution than smoking prevalence, which is expectable. Here, the same interpretation of price coefficient as in the smoking prevalence model. A 10% increase in price results in an 8.09% decrease in smoking demand, which is relatively high compared with other studies.

Figure 4: Cigarette Retail Volume Driver growth decomposition



Looking at the growth decomposition, sadly, the MPOWER index is nowhere to be seen since over the whole sample, there was no change recorded in Italy. Consumer expenditure on alcohol stands out quite a bit. In 2020 the expenditure on alcohol decreased significantly, probably due to covid-19, closed bars, restaurants, etc. The growth in the urban population seems relatively stable. It does not contribute much to the model. Wage represents the disposable income part in the model and does not have drastic changes. Cigarette prices were increasing throughout all the years. However, it does not take up much in terms of affecting changes in the model. Knowing that it will be pretty nice to play around with enforcing bans on tobacco advertising in the government revenue tool since it gives room to improve the index and see how the change would affect government revenue.

4. Possible scenarios

In this section, I will present the possibilities of how to use these models in practice. These results are pretty valuable for governments, which implement taxation policies. In the image below, taxes for forecast years are just repeated values of the available last year (in this case, 2019 taxes). As mentioned before, policymakers can also control the smoking population by adjusting MPOWER indices included in the model. This tool presents three things: smoking population graph, cigarette retail volume graph, and Government revenue, calculated by $CigarettePrice * TotalTax * CigareteVolume$. One possible scenario is increasing MPOWER indices and adjusting the revenue loss by raising cigarette prices through taxes (Scenario 1). Another option is to leave the MPOWER indices alone and only control through taxes (Scenario 2).

4.1 Scenario 1

Table 7: Scenario 1 Taxes

tax	2018	2019	2020	2021	2022	2023	2024
Specific excise (€/1000 sticks)	19.27	20.90	20.90	80.90	84.40	89.90	93.90
Ad-valorem (% RSP)	51.00	50.97	50.97	50.97	50.97	50.97	50.97
VAT (% RSP)	18.03	18.03	18.03	18.03	18.03	18.03	18.03
Total tax (% RSP)	76.45	76.97	76.83	86.39	86.53	86.77	86.96

Table 8: Scenario 1 Prices (€ per stick)

	2018	2019	2020	2021	2022	2023	2024
New Price	0.26	0.26	0.27	0.47	0.48	0.50	0.52
Old Price	0.26	0.26	0.27	0.27	0.28	0.28	0.29

Table 9: Scenario 1 MPOWER

Index	2018	2019	2020	2021	2022	2023	2024
Enforce Bans on Tobacco Advertising	4	4	4	5	5	5	5

Figure 5: Cigarette Demand in Italy Scenario 1

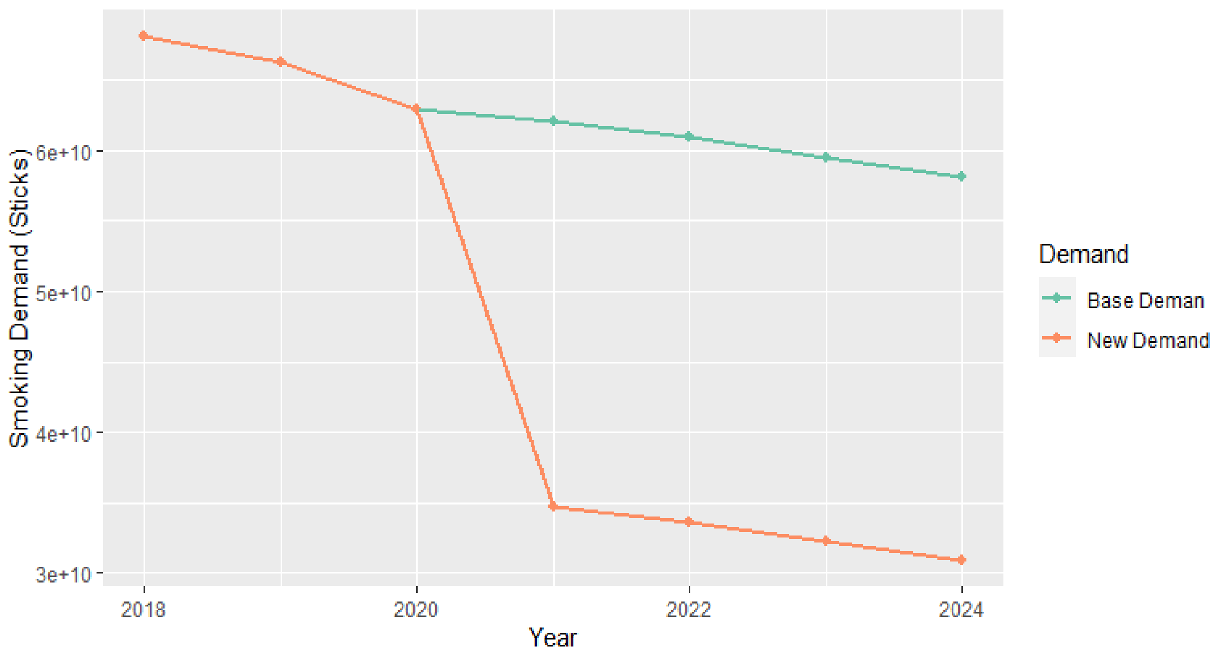


Figure 6: Smoking Prevalence in Italy Scenario 1

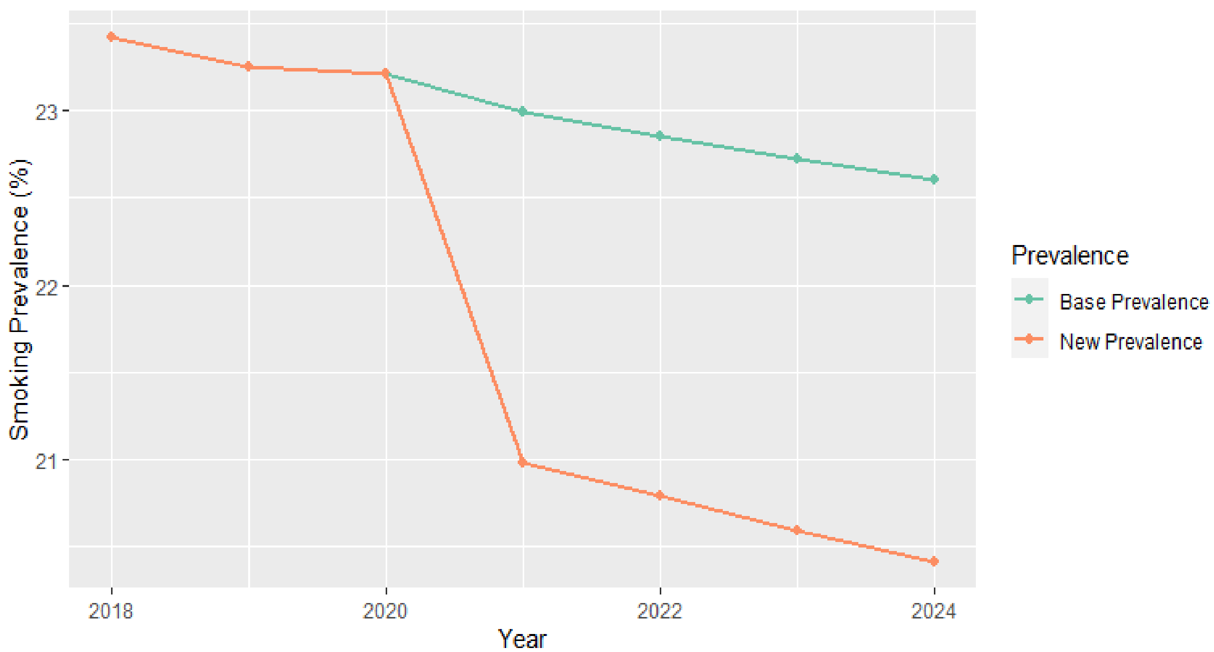
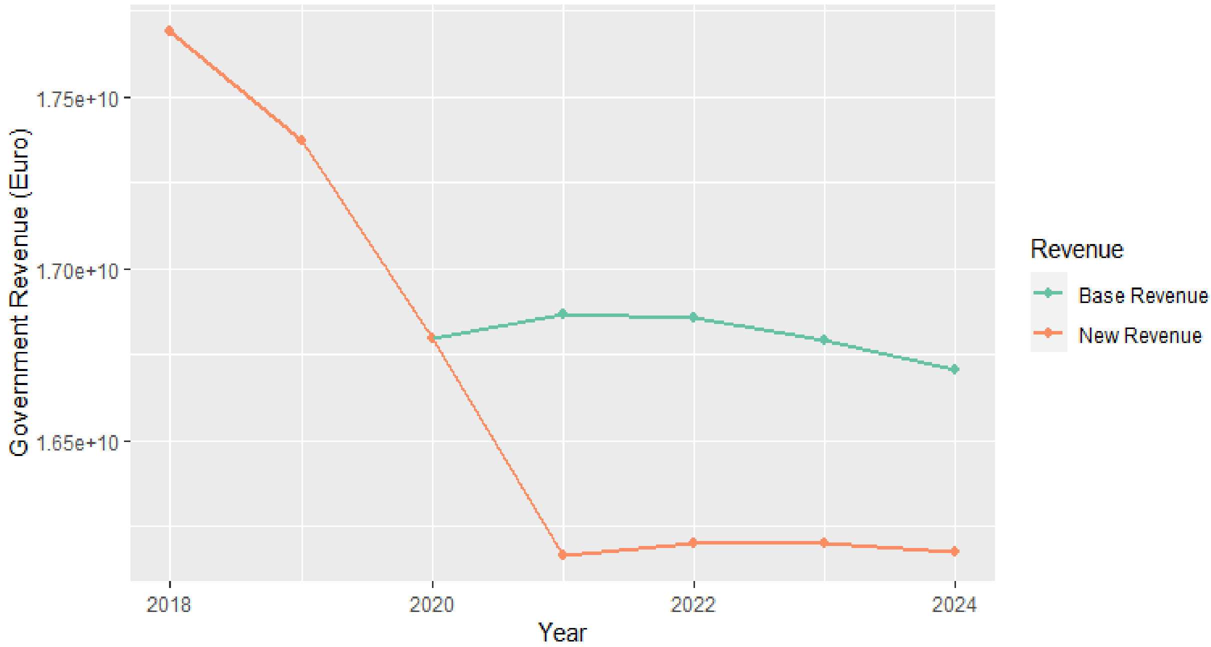


Figure 7: Government Tax Revenue in Italy Scenario 1



Looking at the first scenario, the government revenue that I will be aiming for is constant revenues from 2020 till 2024. Increasing the Enforce Bans on Tobacco Advertising index from four to five would be pretty hard for Italy because most advertisement channels are already forbidden for tobacco products. Mainly indirect advertising is left unbanned. However, if Italy manages to achieve a higher index, consumption will decrease significantly, and the smoking population will also decrease (however, this is not implemented in the model). At this point, government revenues have suffered quite a significant decrease. Therefore, a tax adjustment is implemented. A substantial increase in specific excise tax for 2021, 60€ for 1000 sticks (or 0.06€ for a single stick), would adjust the revenue loss, but still not enough, due to bans on advertising which has a significant impact on the demand for cigarettes. For the following years, additional 3.5€, 5.5€, and 4€ increases are implemented to keep the loss stable and in line with 2021 revenues. Even though per stick taxes have increased only by 0.06€, the total price has risen by 74% due to high ad-valorem and VAT, which are calculated as percentages of Retail Selling Price (or RSP). Therefore, by 2024, the prices will increase (compared to old prices) by 79%. Finally, looking at how much of the final price is an excise tax, based on the previously mentioned WHO benchmark, it is 68%, which is still not high enough. However, it would not be that hard to reach if benchmark levels would be the priority. This scenario would decrease the demand for cigarettes significantly. However, the government would lose quite a lot of revenue due to that.

4.2 Scenario 2

Table 10: Scenario 2 Taxes

tax	2018	2019	2020	2021	2022	2023	2024
Specific excise (€/1000 sticks)	19.28	20.90	20.90	30.90	31.90	34.40	37.90
Ad-valorem (% RSP)	51.00	50.97	50.97	50.97	50.97	50.97	50.97
VAT (% RSP)	18.03	18.03	18.03	18.03	18.03	18.03	18.03
Total tax (% RSP)	76.45	76.97	76.83	79.16	79.22	79.57	80.07

Table 11: Scenario 2 Prices

	2018	2019	2020	2021	2022	2023	2024
New Price	0.26	0.26	0.27	0.30	0.31	0.33	0.34
Old Price	0.26	0.26	0.27	0.27	0.28	0.28	0.29

Table 12: Scenario 2 MPOWER

Index	2018	2019	2020	2021	2022	2023	2024
Enforce Bans on Tobacco Advertising	4	4	4	4	4	4	4

Figure 8: Cigarette Demand in Italy Scenario 2

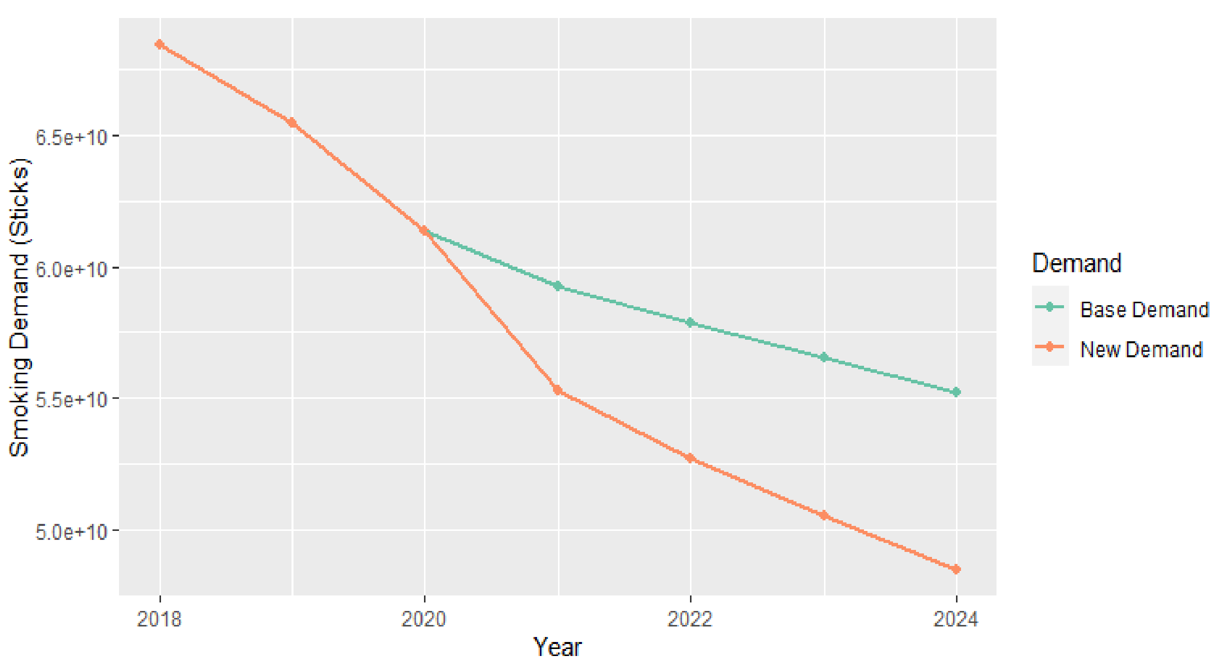


Figure 9: Smoking Prevalence in Italy Scenario 1

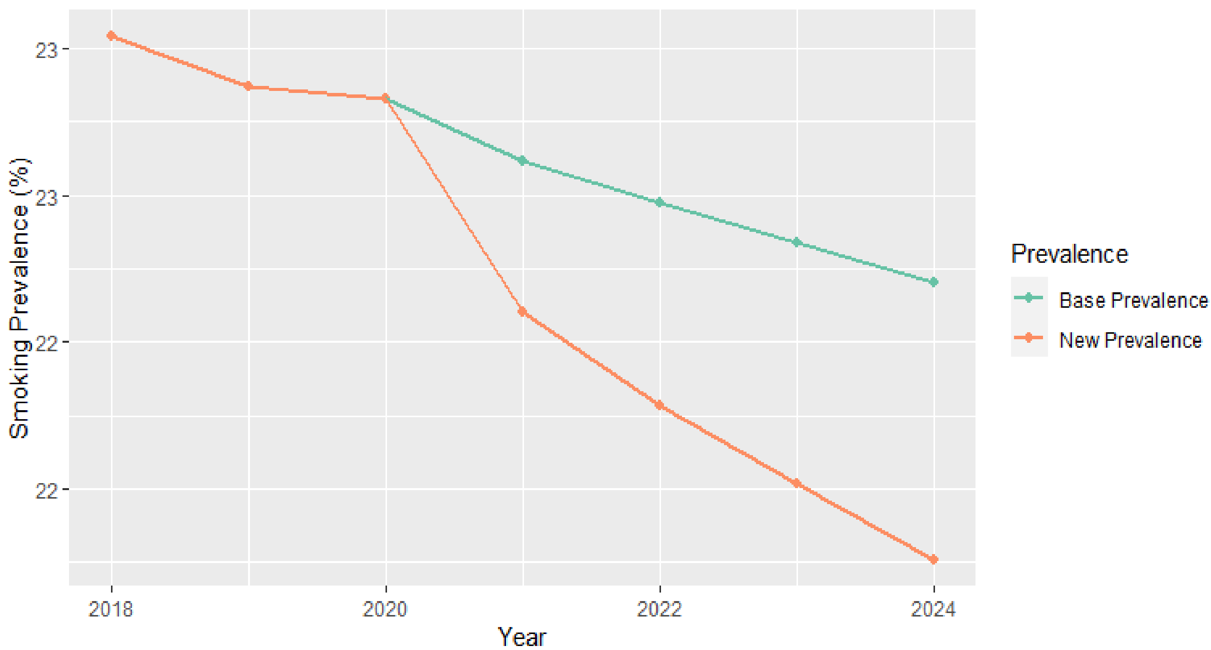
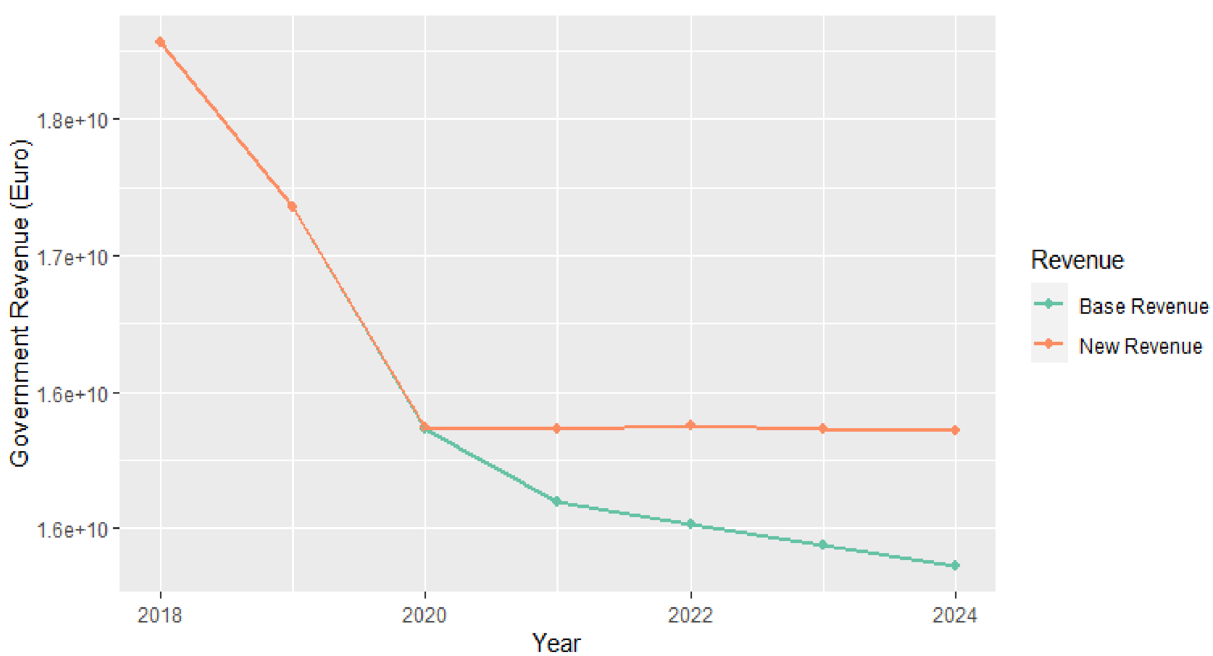


Figure 10: Government Tax Revenue in Italy Scenario 1



When looking at scenario two, MPOWER indices were not changed. The same government revenue levels are being aimed to achieve as in scenario one. Here taxes are not being increased by much. In 2021, only 10€ per 1000 sticks is implemented, which means that the price of a single cigarette increases by 0.01€. However, since Ad-valorem and VAT contribute a lot from the taxation side, the final price of a single stick increases by 0.03€. For the following years, additional 1€, 2.5€, and 2.5€ increases are introduced to keep up with naturally decreasing demand and smoking population. Initially, the price increases by 11%, and by 2024 the price has increased (compared to the old price) by 17%. Comparing to the WHO benchmark, excise taxes make up 60% of RSP, and total taxes make up almost 80% of RSP in 2024.

4.3 Scenario conclusions

Looking at both scenarios final results, the second scenario is more of stability-keeping policy implementation. Helping the government to keep stable revenues while reducing smoking demand. However, the first scenario is more control-based. When additional bans are introduced, smoking demand decreases drastically, and adjusting the government revenues by increasing taxes, reduces demand even more, not compensating for total revenue loss.

However, the models are not perfect for this scenario tool. Optimally, I would be looking for models with the same drivers, or at least MPOWER indices would be the same in both models. Here, introducing additional bans does not reduce smoking prevalence in any way because it is not included in the model. Additionally, another solution would be to introduce soft factors that would account for changes in left-out MPOWER indices.

Conclusion

When comparing my results with previous studies, the price elasticity of demand was relatively higher. However, knowing that the sample size was relatively small and some deviation in the data due to the covid pandemic were present, the results are deemed valid. In Italy, many measures were taken to control tobacco use, and the main way to combat this in the future will be through taxation. Revenue losses will be present, but eventually, everyone is trying to keep the population healthy, and tobacco tax revenues in a completely healthy country will be zero. The demand model produced a -0.809 price elasticity and smoking prevalence -0.212. MPOWER indices were adequate. However, Italy does not have much room to improve them. In the optimal scenarios, if specific excise is increased by 7.5€ for 1000 sticks in 2021, total cigarette demand reduces by 6.66%, smoking prevalence by 0.41%, and total government revenues increase by 1.64%. In the case where MPOWER indices were improver, adjusting government revenue by increasing specific excise by 50€ per 1000 sticks in 2021 gave a 36.24% reduction in demand, 2.16% reduction in smoking prevalence, and 1.6% increase in government revenue.

References

1. The Tobacco Atlas. Italy summary: https://tobaccoatlas.org/country/italy/
2. Yeh, CY., Schafferer, C., Lee, JM. *et al.* The effects of a rise in cigarette price on cigarette consumption, tobacco taxation revenues, and of smoking-related deaths in 28 EU countries-- applying threshold regression modelling. *BMC Public Health* **17,** 676 (2017). https://doi.org/10.1186/s12889-017-4685-x
3. Lien Nguyen, Hunnar Rosenqvist, Markku Pekurinen. Demand for Tobacco in Europe, an Econometric Analysis of 11 countries for the PPACTE Project. *National Institute for health and welfare*, (2012) https://www.julkari.fi/bitstream/handle/10024/90864/URN_ISBN_978-952-245-594-9.pdf?sequence=1
4. Dr. Esteve Fernández. Economic Aspects of Smoking In Europe. *A European Project* from 15 September 2001 – 15 September 2002. https://ec.europa.eu/health/ph_projects/2001/cancer/fp_cancer_2001_a1_05.pdf
5. Levy, D., Gallus, S., Blackman, K. *et al.* Italy SimSmoke: the effect of tobacco control policies on smoking prevalence and smoking attributable deaths in Italy. *BMC Public Health* **12,** 709 (2012). https://doi.org/10.1186/1471-2458-12-709
6. Fuertes, E., Marcon, A., Potts, L. *et al.* Health impact assessment to predict the impact of tobacco price increases on COPD burden in Italy, England and Sweden. *Sci Rep* **11,** 2311 (2021). https://doi.org/10.1038/s41598-021-81876-3
7. Indahl UG, Næs T, Liland KH. A similarity index for comparing coupled matrices. Journal of Chemometrics. 2018;e3049. https://doi.org/10.1002/cem.3049
8. Robert P, Escoufier Y. A unifying tool for linear multivariate statistical methods: the RV-coefficient. Appl Stat. 1976;25(3):257-265 https://www.jstor.org/stable/2347233?seq=1
9. Smilde AK, Kiers HA, Bijlsma S, Rubingh CM, van Erk MJ. Matrix correlations for high-dimensional data: the modified RV-coefficient. Bioinformatics. 2009 Feb 1;25(3):401-5. doi: 10.1093/bioinformatics/btn634. Epub 2008 Dec 10. PMID: 19073588.
10. Additional information about MPOWER indices: https://www.who.int/tobacco/mpower/en/

EFFECTIVENESS OF TOBACCO TAXATION ON SMOKING PREVELENCE AND CONSUMPTION: A CASE OF ITALY

LUKAS DRAZDYS

Bachelor Thesis

Quantitative Economics programme

Faculty of Economics and Business Administration of Vilnius University

Supervisor – Dr. Andrius Kažukauskas

Vilnius, 2021

Summary

48 pages, 10 pictures, 12 tables, 10 references.

The thesis goal was to identify how tobacco taxes and prices affect the smoking prevalence and their demand. The paper consists of three main parts: literature analysis, model creation and their result, government revenue tool creation.

Literature analysis was conducted to see what kind of results I should expect in created models. To see what possibly drives tobacco consumption and what approach is better.

Model creation was based on fixed-effect panel models using a specially made model selection algorithm, which filtered out a long list of drivers by testing them according to formed expectations of driver performance. The main purpose of using an algorithm instead of forming the models by hand is to test more drivers and better see what possible causes higher or lower tobacco consumption. The results provided a higher price elasticity of demand than previous studies. However, the price elasticity of smoking prevalence was in line with previous researches.

A government revenue tool was then built, using those created models, so it would be a lot easier to see how an increase in tax affects the price, smoking prevalence, cigarette demand, and government revenue. This can be a handy tool for governments to see what could be an optimal policy to implement.

TABAKO APMOKESTINIMO VEIKSMINGUMAS RŪKANČIAI POPULIACIJAI IR VARTOJIMUI: ITALIJOS ATVEJIS

LUKAS DRAZDYS

Bakalauro darbas

Kiekybinės ekonomikos programa

Vilniaus universitetas Ekonomikos ir Verslo Administravimo Fakultetas

Vadovas – Dr. Andrius Kažukauskas

Vilnius, 2021

Santrauka

48 puslapiai, 10 paveikslėliai, 12 lentelės, 10 šaltiniai.

Darbo tikslas buvo nustatyti, kaip tabako mokesčiai ir kainos daro įtaką jo paklausai bei rūkančiajai populiacijai. Darbą sudaro trys pagrindinės dalys: literatūros analizė, modelių kūrimas ir jų rezultatai, vyriausybės pajamų įrankio kūrimas.

Literatūros analizė buvo atlikta siekiant sužinoti, kokių rezultatų galima tikėtis kuriant modelius. Taip pat sužinoti, kas galbūt skatina tabako vartojimą bei kokią metodologiją reikėtų taikyti ieškant rezultatų.

Modeliai buvo kuriami naudojant fiksuotų efektų panelinius modelius, naudojant specialiai sukurtą algoritmą, kuris išfiltruoja didelį sarašą modelio kintamųjų, juos testuojant pagal susidarytus lūkesčius. Pagrindinė priežastis, kodėl buvo naudojamas algoritmas, tai išbandyti didesnį kiekį kintamųjų, ir geriau suprasti kaip kiekvienas iš jų veikia modelį. Rezultatai suteikė didesnį kainos elastingumą kainai, nei ankstesni tyrimai. Tačiau, rūkančiosios populiacijos kainos elastingumas atitiko ankstesnių tyrimų rezultatus.

Turint modelius, buvo sukurtas vyriausybės pajamų įrankis, kad būtų daug lengviau suprasti, kaip mokesčių padidėjimas veikia kainą, rūkančiąją populiaciją, cigarečių paklausą ir vyriausybės pajamas. Tai yra labai naudinga priemonė vyriausybėms, norint sužinoti, kokį mokesčio padidinimą reikėtų įvesti.

Annex

Annex 1: Driver List and Expected Signs with definitions

Industry	Category	Sign	Definition	Source
Alcohol	Alcoholic Drinks Price	-1	Alcoholic drinks is the aggregation of beer, wine, spirits, cider/perry and RTDs.	Passport
Economy	Real GDP Growth	1	The number reached by valuing all the productive activity within the country at a specific year's prices. When economic activity of two or more time periods is valued at the same year's prices, the resulting figure allows comparison of purchasing power over time, since the effects of inflation have been removed by maintaining constant prices.	Passport
Economy	GDP	1	Gross domestic product is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources.	Passport
Employment	Employment Rate	1	Employed population aged 15-64 as a percentage of working age (15-64 years) population. In Passport Cities data, employment rate refers to total employed population as percentage of working age (15-64) population.	Passport
Employment	Male Employment Rate	1	Male employed population aged 15-64 as a percentage of working age (15-64 years) male population.	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Employment	Female Employment Rate	1	Female employed population aged 15-64 as a percentage of working age (15-64 years) female population.	Passport
Employment	Unemployment Rate	0	Unemployment rate represents unemployed population as a percentage of the economically active population, also known as the labour force (the total number of people employed plus unemployed). The ILO international standard definition of unemployment is based on the following three criteria which should be satisfied simultaneously: “without work”, “currently available for work” and “seeking work”.	Passport
Employment	Male Unemployment Rate	0	Male unemployment rate represents unemployed male population as a percentage of the male economically active population, also known as the labour force (the total number of people employed plus unemployed). The ILO international standard definition of unemployment is based on the following three criteria which should be satisfied simultaneously: “without work”, “currently available for work” and “seeking work”.	Passport
Employment	Female Unemployment Rate	0	Female unemployment rate represents unemployed female population as a percentage of the female economically active population, also known as the labour force (the total number of people employed plus unemployed). The ILO international standard definition of unemployment is based on the following three criteria which should be satisfied simultaneously: “without work”, “currently available for work” and “seeking work”.	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Expenditure	Consumer Expenditure on Fruit	-1	Fresh, chilled or frozen fruit, dried fruit, fruit peel, fruit kernels, nuts and edible seeds, preserved fruit and fruit-based products, melons and water melons.	Passport
Expenditure	Consumer Expenditure on Alcoholic Beverages and Tobacco	1	The alcoholic beverages purchased for consumption at home and all purchases of tobacco (cigarettes, cigars, cigarillos, pipe tobacco, hand rolling tobacco, snuff and smoking accessories) by households, including purchases of tobacco in restaurants, cafés, bars, service stations, etc.	Passport
Expenditure	Consumer Expenditure on Alcoholic Drinks	1	The alcoholic beverages classified here are those purchased for consumption at home. The group excludes alcoholic beverages sold for immediate consumption away from the home by hotels, restaurants, cafés, bars, kiosks, street vendors, automatic vending machines, etc. (consumer expenditure on catering). The beverages classified here include low or non-alcoholic beverages, which are generally alcoholic such as non-alcoholic beer.	Passport
Expenditure	Consumer Expenditure on Water and Miscellaneous Domestic Services	-1	Water supply, refuse, sewage collection and disposal, other services relating to the dwelling n.e.c. Associated expenditure such as hire of meters, reading of meters, standing charges, co-proprietor charges for caretaking, gardening, stairwell cleaning, heating and lighting, maintenance of lifts and refuse disposal chutes, etc. in multi-occupied buildings, security services, snow removal and chimney sweeping.	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Expenditure	Government Expenditure on Education	-1	The data refers to all non-repayable payments by General Government - whether capital or current, requited or not. General Government expenditure on education includes following items. Capital expenditure includes expenditure for construction, renovation and major repairs of buildings and the purchase of heavy equipment or vehicles. Current expenditure includes expenditure for goods and services consumed within the current year and which would need to be renewed if there were a need for prolongation the following year. It includes expenditure on: staff salaries and benefits; contracted or purchased services; other resources including books and teaching materials; welfare services; and other current expenditure such as furniture and equipment, minors repairs, fuel, telecommunications, travel, insurance and rents. If General Government is financing any private education institution in any ways (usually – subsidies), then this expenditure is included too.	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Expenditure	Government Expenditure on Health	-1	The data refers to all non-repayable payments by General Government - whether capital or current, requited or not. General Government expenditure on health includes: medical products, appliances and equipment (such as pharmaceutical products, other medical products, and therapeutic appliances and equipment), outpatient services (such as general medical services, specialised medical services, dental services, and paramedical services), hospital services (such as general hospital services, specialised hospital services, medical and maternity centre services, and nursing and convalescent home services), public health services, R & D on health, health n.e.c.	Passport
Expenditure	Consumer Expenditure on Electricity Gas and Other Fuels	-1	Expenditure on electricity, gas, heat energy, liquid and solid fuels.	Passport
Expenditure	Consumer Expenditure on Education	-1	Expenditure on pre-primary and primary education, secondary education, post-secondary non-tertiary education, tertiary education, education not definable by level. This division covers educational services only.	Passport
Household	Average Number of Children Per Household	1	Average number of children aged 0 – 17 per household.	Passport
Household	Average Size of Urban Household	0	Refers to the average number of people per household in an urban area.	Passport
Household	Average Size of Rural Household	0	Refers to the average number of people per household in a rural area.	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Household	Average Household Size	0	Refers to average number of people per household. A household is a small group of persons who share the same living accommodation, who pool some, or all, of their income and wealth and who consume certain types of goods and services collectively, mainly housing and food. The members of a household thus defined are not necessarily related by blood or marriage. Resident domestic servants are included.	Passport
Household	Number of Households	0	A household is a small group of persons who share the same living accommodation, who pool some, or all, of their income and wealth and who consume certain types of goods and services collectively, mainly housing and food. The members of a household thus defined are not necessarily related by blood or marriage. Resident domestic servants are included	Passport
Household	Households with 1 Child	0	Refers to households with one child under 18 years old.	Passport
Household	Households with 2 Children	0	Refers to households with one child under 18 years old.	Passport
Household	Households with 3 Children	0	Refers to households with one child under 18 years old.	Passport
Household	Households with 4 and more Children	0	Refers to households with one child under 18 years old.	Passport
Income	Minimum Wage per Hour	1	Minimum wage is a minimum level of gross amount, that is before deduction of income tax and social security contributions and which is established by law for work performed per month. The data refers to a minimum wage which is valid on the 1st of July.	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Income	Wage per Hour in Manufacturing	1	Wage per hour in manufacturing refers to average wage per worker in manufacturing as a whole, according to the International Standard Industrial Classification of all economic activities (ISIC). The series cover workers of both sexes, irrespective of age. Statistics relates to gross remuneration in cash and in kind paid to employees, as a rule at regular intervals, for time worked or work done together with remuneration for time not worked, such as annual vacation, other type of paid leave or holidays.	Passport
Income	Wage per Hour	1	Wage per hour refers to average wage per worker. The series cover workers of both sexes, irrespective of age.	Passport
Income	Disposable Income	1	This is gross income minus social security contributions and income taxes.	Passport
Life Standards	Percent of Population Aged 15 Plus with Higher Education	-1	Refer to the % of population aged 15+ with higher education. Higher education refers to ISCED levels 5, 6 and 7, provided at universities, teachers' colleges, and higher professional schools, and requiring, as a minimum condition of admission, the successful completion of secondary education or evidence of the attainment of an equivalent level of knowledge.	Passport

Table 2: Driver List and Expected Signs (Continued)

Industry	Category	Sign	Definiton	Source
MPOWER	Monitor	-1	The implementation status of the Monitoring measure was classified by grouping countries into four groups. The groups for this indicator are: 1 = No known data or no recent* data or data that are not both recent* and representative** 2 = Recent* and representative** data for either adults or youth 3 = Recent* and representative** data for both adults and youth 4 = Recent, *representative* ***and periodic*** data for both adults and youth * Recent means? in the data year? or the 5 years ?previous to the data year. ** Survey sample representative of the national population. *** Survey repeated at least every five years. Survey is considered an "adult survey" if it is a household survey sampled from the general population with respondent ages not limited to those under 15.	World Health Organization
MPOWER	Offer Help to Quit Tobacco Use	-1	Information from countries on the availability and non-availability of particular tobacco cessation aids is assessed to determine the comparative level of assistance countries provide to help tobacco users quit. The groupings for this indicator are: 1 = Data not reported 2 = None 3 = NRT* and/or some cessation services** (neither cost-covered) 4 = NRT* and/or some cessation services** (at least one of which is cost-covered) 5 = National quit line, and both NRT* and some cessation services** cost-covered * Nicotine replacement therapy. ** Smoking cessation support available in any of the following places: health clinics or other primary care facilities, hospitals, office of a health professional, the community.	World Health Organization

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
MPOWER	Protect from Tobacco Smoke	-1	Country's legislation is assessed to determine whether smoke-free laws exist in each of the following places at either the national or subnational level: • health-care facilities; • educational facilities other than universities; • universities; • government facilities; • indoor offices; • restaurants; • pubs and bars; • public transport. The implementation status of the Smoke-free environments measure was classified by grouping countries into five groups. The groupings for this indicator are: 1 = data not reported/not categorized* 2 = Up to two public places completely smoke-free 3 = Three to five public places completely smoke-free 4 = Six to seven public places completely smoke-free 6 = All public places completely smoke-free (or at least 90% of the population covered by complete subnational smoke-free legislation)	World Health Organization
MPOWER	Warn About the Dangers of Tobacco	-1	Country's legislation is assessed to determine whether health warnings with specific criteria are mandated. The groupings for this indicator are: 1 = data not reported 2 = No warning or warning covering <30% of pack surface 3 = ?30%* but no pictures or pictograms and/or other appropriate characteristics** 4 = 31%–49%* including pictures or pictograms and other appropriate characteristics** 5 = ?50%* including pictures or pictograms and appropriate characteristics** * average of the front and back of the cigarette pack. ** • Specific health warnings mandated; • appearing on individual packages as well as on any outside packaging and labelling used in retail sale; • describing specific harmful effects of tobacco use on health; • are large, clear, visible and legible (e.g. specific colours and font style and sizes are mandated); • rotate; • written in (all) principal language(s) of the country	World Health Organization

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
MPOWER	Enforce Bans on Tobacco Advertising	-1	Country's legislation is assessed to determine whether all or any forms of tobacco advertising promotion and sposorship are banned. The groupings for this indicator are: 1 = data not reported 2 = Complete absence of ban, or ban that does not cover national television (TV), radio and print media 3 = Ban on national TV, radio and print media only 4 = Ban on national TV, radio and print media as well as on some but not all other forms of direct* and/or indirect** advertising 5 = Ban on all forms of direct* and indirect** advertising. * Direct advertising bans: • national television and radio; • local magazines and newspapers; • billboards and outdoor advertising; • point of sale. ** Indirect advertising bans: • free distribution of tobacco products in the mail or through other means; • promotional discounts; • non-tobacco products identified with tobacco brand names (brand extension); • brand names of non-tobacco products used for tobacco products; • appearance of tobacco products in television and/or films; • sponsored events.	World Health Organization
Population	Dependency Ratio	1	Indicates the percentage of population aged 0 – 14 and persons older than 65 per persons aged 15 – 64.	Passport
Population	Death Rates	-1	The death rate is the annual number of deaths per 1,000 population. Death is defined as permanent disappearance of all evidence of life at any time after live birth has taken place (post-natal cessation of vital functions without capability of resuscitation). This definition therefore excludes foetal deaths.	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Population	Deaths	-1	Permanent disappearance of all evidence of life at any time after live birth has taken place (post-natal cessation of vital functions without capability of resuscitation). This definition therefore excludes foetal deaths.	Passport
Population	Life Expectancy at Birth	0	Indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.	Passport
Population	Net Migration Rate	0	The difference between the number of immigrants into and emigrants from the area during the year.	Passport
Population	Generation Z	-1	Generation Z (also known as the Internet Generation) is a common name for the group of people born between 1995 and 2009.	Passport
Population	Millennials	0	Millennials refer to the population group born from 1980 to 1994. It is also known as the Generation Y or the Echo Boomers - the demographic cohort following Generation X.	Passport
Population	Generation X	0	Generation X is the generation born after the Western post–World War II baby boom ended between the years 1965 and 1979.	Passport
Population	Teens Aged 13-17	0	This age grouping covers the ages 13-17.	Passport
Population	Young Adults Aged 18-29	0	This grouping covers the ages 18-29.	Passport
Population	Middle Youth Aged 30-44	0	This grouping covers the ages 30-44.	Passport
Population	Mid-lifers Aged 45-64	0	This grouping covers the ages 45-64.	Passport
Population	Later Lifers Aged 65-79	0	This grouping covers the ages 65-79.	Passport
Population	Seniors Aged 80 Plus	0	This group covers elderly people aged 80 plus.	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Population	Urban Population	-1	Population of areas defined as urban in each country and reported to the United Nations. The infrastructure, types of houses and dwellings in agglomerations, density, landscapes, etc., are very different between the countries, hence, naturally, it is impossible to have a uniform definition.	Passport
Tobacco	Cigarette Price	-1	RETAIL SALES OF DUTY PAID CIGARETTES The definition of cigarettes for the purposes of this study is duty-paid, machine manufactured white-stick products. This does not exclude brands of cigarettes that do not use white paper but it is designed to exclude the volume of non-machine manufactured products such as bidis/beedis (India) and papirosy (Russia), and other smoking products made with tobacco but that either do not resemble cigarettes as recognised in the US or Europe, or those that are not machine manufactured.	Passport
Tobacco	Cigarillos Price	1	Cigarillos are defined as miniature cigars weighing less than 3 grams each, with a ring gauge of <29. The distinction between small cigars and cigarillos is relatively blurry in the tobacco industry, and there are many cigarillos marketed as small (or 'little') cigars.	Passport
Tobacco	Fine Cut Tobacco Price	1	Fine Cut tobacco includes tobacco sold in packaged format for use in roll-your-own (RYO) and make-your-own (MYO) cigarettes	Passport

Annex 1: Driver List and Expected Signs with definitions (Continued)

Industry	Category	Sign	Definiton	Source
Tobacco	Smoking Tobacco Price	1	Smoking tobacco consists of cut tobacco sold in packaged format for smoking either in pipes or for use in roll-your-own (RYO) cigarettes. This excludes smokeless tobacco, such as snuff (or ‘snus’) and chewing tobacco.	Passport
Tobacco	Cigars Price	1	This category is the aggregation of large, standard and small cigars only. This category excludes cigarillos. Company and brand shares are available at this level for the combined market of large, medium and small cigars, with subsector level splits available only for market sizes (eg volume and value sales). NB Euromonitor uses ring gauge to distinguish cigarillos from small cigars.	Passport
Tobacco	Pipe Tobacco Price	1	Pipe tobacco includes cut tobacco sold in packaged format for smoking in pipes. It also includes water pipe tobacco of the type consumed in the Middle East, known as ‘’shisha’’.	Passport
Economy	Human Development Index	-1	The Human Development Index (HDI) is an index used to rank countries by level of “human development”. The HDI provides a composite measure of three dimensions of human development: living a long and healthy life (measured by life expectancy), being educated (measured by adult literacy and gross enrolment in education) and having a decent standard of living (measured by purchasing power parity, PPP, income). The HDI sets a minimum and a maximum for each dimension, called goalposts, and then shows where each country stands in relation to these goalposts, expressed as a value between 0 and 1, where 0 shows the lowest HDI value and 1 shows the highest. The scores for the three HDI dimension indices are then aggregated into a composite index using geometric mean.	Passport