Vol. V, Issue 1
February 2015

The Determinants of Tobacco Consumption: Evidence from Spain

Authors: **Marta Pascual Sáez,** Department of Economics, Faculty of Economics and Business, University of Cantabria, Spain, marta.pascual@unican.es; **Noelia González-Prieto,** Department of Economics. Faculty of Economics and Business, University of Cantabria, Spain, gprieton@unican.es; **David Cantarero-Prieto,** Department of Economics. Faculty of Economics and Business University of Cantabria, Spain, david.cantarero@unican.es

The aim of this paper is to study the determinants of tobacco consumption in Spain. First, the relationship between tobacco consumption and Gross Domestic Product (GDP) in Spain is analyzed.

The non-stationary and cointegration properties of these variables are studied. The findings suggest that tobacco mortality is not cointegrated with GDP.

Afterwards, we model the probability of an individual being smoker in Spain as a function of a range of socio-economic characteristics, including individual's gender, age, marital status, education, etc., using data from the European Community Household Panel (ECHP).

The results show that men with lower educational background and unemployed are more likely to smoke.

Key words: *Unit Root, Cointegration, Smoking, ECHP, GDP, Spain.*

JEL Classification*: I18, H51*

Introduction

During the last years, population health has been considered as a fundamental issue in all countries and one of the most important indicators of life quality. In this way, policy makers have an increased interest in social inequalities in health and in those characteristics of individuals that are related to health. Smoking is one of the largest causes of death and disease in the European Union and policy makers are concerned about it. More than 500,000 Europeans die every year because they smoke and over 13 million more are suffering from a serious chronic disease as a result of smoking. Also, passive smoking has important effects on health (cancer, heart disease, asthma attacks, coughing, eye irritation, etc).

The World Health Organization (WHO) forecasts that between 2000 and 2030 the number of smokers will rise from 1.2 billion to 1.6 billion and the annual number of deaths will increase from 4.9 million to 10 million (WHO, 2004). WHO (2012) estimates that death rates (per 100,00) attributable to tabacco in 2004 is 1,444 for ages 30 and over. Furthermore, in Spain, 23.95% of the population, aged 15 years old and over, are daily smokers (National Health Survey 2011-2012). Thus, smoking is associated with higher morbidity rates, mortality and health expenditure. Even tobacco is associated with lower socio-economic groups in which smoking is more prevalent (see Ross 2004, European Network for Smoking Prevention project report 2004). In this sense, Regidor et al. (2001) evaluates smoking trends among Spanish men and women by social class between 1987 and 1997 concluding that among men ages 25 years and older the prevalence of smoking in both the manual and the non-manual social class decreased in all age groups. In contrast, among women the prevalence of smoking increased in both social classes.

As well, tobacco smoke is an important serious environmental hazard. On one hand, Rovira et al. (2000) explores the public attitudes toward environmental tobacco risks in Spain comparing the perception of passive smoking risks to the individual perceptions of risk to themselves from smoking. On the other hand, different authors such as Waldron (1991), Graham (1996) and Regidor et al. (2001) have noted that the spread of the smoking epidemic in the developed countries follows roughly four stages. During the first stage, smoking is infrequent in the population and primarily affects men in the highest socioeconomic strata. Secondly, the prevalence of

smoking among men increases considerably in all socioeconomic groups. In the third stage, men begin to give up smoking while the prevalence among women increases. Finally, during the fourth stage, the prevalence of smoking slowly decreases among both men and women.

Pharmaceuticals are becoming an important issue of healthcare in developed countries not only from a clinical point of view but also financially. In fact, many countries are very concerned in how much they spend on health and the rate at which it grows. "Health at a Glance 2013" (OECD, 2013) provides the latest comparable data on different aspects of the performance of health systems in OECD countries. However, it is important to point out the following data: "On average across the OECD, per capita health spending grew by 4.1% annually in real terms over 2000-2009 but this slowed to 0.2% in 2009-10 and 2010-11 as many countries reduced health spending to help cut budget deficits and government debt, especially in Europe".

Taking into account all these considerations, this paper attempts to contribute to the literature on the determinants of tobacco consumption. The rest of the paper is structured as follows. Section 2 describes the methodological decisions adopted. And finally, Section 3 describes the data and the empirical results. Finally, concluding remarks are shown in last section.

Theoretical Framework

In order to assess the main determinants of tobacco consumption, we specify our baseline model, in which we analyze different time series. In particular, we study those variables which could be related with tobacco consumption. As it is well known, a time series can be defined as a collection of random variables ordered in time (t). From a statistical point of view, a stationary process is one whose statistical properties do not change over time. Thus, if a time series is stationary, the mean and the variance of the stochastic process do not depend on t and the auto covariance between Y_t and Y_{t+k} only depends on the lag k. So, a stationary series would contain no trend or seasonal variation. Furthermore, a time series is said to be integrated of order d, denoted by $I(d)$, if you have to difference it d times to obtain a stationary process. Thus, we say a time series Y_t is integrated of

order 1, $I(1)$, if Y_t is not stationary but the first difference, Y_t-Y_{t-1}, is stationary and invertible (see Maddala, 1992).

The most common test in economic literature for unit roots are Augmented Dickey and Fuller unit root test (ADF, 1979) and Phillips-Perron (1978). The ADF test involves estimating the following regression:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{k} \delta_i \Delta y_{t-1} + \varepsilon_t \quad (1)$$

which includes a trend t, a constant term α, as well as lagged versions of the series, and where y_t is the variable of interest. The null hypothesis for this test is H_0: $\hat{\gamma} = 0$.

On the other hand, Phillips-Perron test (PP) is a unit root test used to test the null hypothesis that a time series is integrated of order 1. PP is a non-parametric test based on asymptotic theory which works well in large samples. This test estimates autocorrelations in the error process, rather than white noise errors. For this reason, this test is more generally applicable. In fact, Davidson and MacKinnon (2006) conclude that Phillips-Perron test could perform worse in finite samples than the ADF test.

During the last decades, cointegration techniques (Granger, 1981) have generated very much interest among economists (see Johansen, 1991). Two variables x_t and y_t are said to be cointegrated if there exists a parameter α such that $y_t = \alpha x_t + u_t$ is a stationary process. So, it is necessary all the variables to have the same integration order. Otherwise, variables would not have direct causal connection.

In Spain from 1st January 2006, new laws ban smoking in public places including food shops and all workplaces. Smoking has been demonstrated to be responsible for many diseases such as different types of cancer, heart attacks, etc. However, it is difficult to give up smoking as nicotine is a highly addictive substance (Suranovic et al. 1999, Bask and Melkersson 2003). Also, there exists evidence that information about diseases due to smoking does not discourage smoking initiation (Sato and Ohkusa, 2003).

So, we are interested not only in the macro-analysis but also on those personal characteristics of smokers versus non-smokers. Therefore, our dependent variable in the statistical model is a dichotomy variable which takes a value of 1 if the individual smokes daily or occasionally and zero otherwise. Thus, factors such as age, education, marital status, etc., and

some economic data could be relevant in explaining whether an individual smokes.

In this way, the respondent is smoker (Y=1) or not (Y=0) in the corresponding period. A set of factors, such as age, marital status, education, etc., gathered in a vector x explain this fact so that:

$$\text{Prob}(Y=1) = F(x,\beta), \qquad (2)$$
$$\text{Prob}(Y=0) = 1 - F(x,\beta).$$

The set of parameters β reflects the impact of changes in x on the probability. In order to estimate this equation, a nonlinear specification of $F(.)$ can prevent logical inconsistency and the possibility of predicted probabilities outside the range [0,1]. The most common nonlinear parametric specifications are logit and probit models which have been analysed. So, we will use a latent variable interpretation (Jones 2001, Greene, 2003). Let

$$y = 1 \quad if \;\; y_i^* > 0 \qquad (3)$$
$$y = 0 \quad if \;\; y_i^* \le 0$$

where

$$y^* = x'\beta + \varepsilon, \quad (4)$$

and if the distribution is symmetric, such as the normal and logistic, then:

$$\Pr(y=1 \mid x) = \Pr(y_i^* > 0 \mid x) = \Pr(\varepsilon < x'\beta \mid x) = F(x'\beta). \qquad (5)$$

If we assume that ε has a standard normal distribution, we obtain the probit model, while assuming a standard logistic distribution, we obtain the logit model. These models are usually estimated by maximum likelihood estimation and the log-likelihood for a sample of independent observations is:

$$\ln L = \sum_{i=1} \left\{ y_i \ln F(x_i'\beta) + (1-y_i) \ln\left[1 - F(x_i'\beta)\right] \right\}. \qquad (6)$$

Data Description and Results

The data used in this paper are obtained from the Organization for Economic Co-operation and Development (OECD) Health Data. This dataset contains annual data from 1960 to 2012 about health status, health

care resources, expenditure on health, social protection and other economic and demographic references for OECD countries. This information allows us to compare the results and main statistics about health for different countries. In this study we have used information available from 1980 to 2012 about total expenditure on health (EXPENDITURE) and Gross domestic product (GDP), both of them, per capita purchasing power parity. We have also included variables for life expectancy at birth (in years) by gender, LE_FEMALE and LE_MALE and tobacco mortality measured as "Malignant neoplasms of trachea, bronchus and lung (deaths per 100.000 population)". In addition, we have included in this analysis the percentage of population which is sixty five years old and over (P65). The definition of each variable used is given in Table 1.

Table 1: Variable definitions

Name	Definition
LE_FEMALE	Life expectancy females at birth - Years
LE_MALE	Life expectancy males at birth - Years
EXPENDITURE	Total expenditure on health - /capita, US$ PPP
GDP	Gross domestic product - /capita, US$ PPP
TOBACCO1	Tobacco consumption (% of population aged 15+ who are daily smokers)
TOBACCO2	Tobacco consumption (cigarettes per smoker per day)
TOB_MORTALITY	Malignant neoplasms of trachea, bronchus and lung (deaths per 100.000 population)
P65	Population: 65 and over - % total population

Source: Authors´ elaboration

The results of ADF and PP unit root tests reported in Table 2 suggest that all the variables are not integrated of order one. In fact, the order of integration for each variable is not the same. Variables are either I(2), I(1) or I(0) so it is not possible to apply cointegration techniques in most of the cases. With respect with tobacco consumption, either % of population aged 15+ who are daily smokers or cigarettes per smoker per day, it is not possible to deep in the analysis because of insufficient number of observations.

Table 2: Results of ADF and PP unit root test (Spain)

Variables	ADF	PP	Order of Integration
LEFEMALE	-6.6827	-6.6516	I(1)
LEMALE	-9.1150	-23.5738	I(1)
EXPENDITURE	-3.8860	-4.3180	I(2)
GDP*	-2.9558	-2.7528	I(2)
TOBACCO1	-	-	-
TOBACCO2	-	-	-
TOB_MORTALITY	-4.5499	-4.9920	I(0)
P65	-4.0634	-4.8469	I(2)

NOTE: All the variables are statistically significant at the conventional level (that is, 1, 5 and 10%). GDP is statistically significant at 10%.
Source: Authors´elaboration.

Once we have test that life expectancy, GDP, tobacco mortality and health expenditure are not cointegrated with the same order, we will focus our attention on the micro-data contained in the European Community Household Panel (ECHP). This survey contains data on individuals and households for the European Union countries with eight waves available (1994-2001). The main advantage is that information is homogeneous among countries since the questionnaire is similar across them. In fact, it is the first fixed and harmonized panel for studying socio-economic factors of the households and individuals inside the European Union.

This representative survey of households of different European Union countries was carried out for the first time in 1994 and were interviewed 60,500 households (approximately 170,000 individuals). This source of data is coordinated by the Statistical Office of the European Communities (EUROSTAT) and it includes rich new information about income, education, employment, etc. However, only from 1998 to 2001 there is available information about smoking in Spain.

We will focus our analysis on a simple question "Do you smoke or did you ever smoke?" and it takes the values "1" (smoke daily), "2" (occasionally), "3" (not now, used to smoke daily), "4" (not now, used to smoke occasionally) and "5" (never smoked). Table 3 reports frequencies for the response to the question "Do you smoke or did you ever smoke?" for each year. It can be noted that from 1998 to 2001 around fifty percent of the

population has smoked (daily or occasionally). Nevertheless, as we are interested in combining individuals' characteristics with households' income, we will use the information corresponding to 2000.

Table 3: Frequencies for the response to the question: "Do you smoke or did you ever smoke?". Country: Spain

	1998	**1999**	**2000**	**2001**
Smoke daily	29.81	28.21	27.35	27.83
Smoke occasionally	4.94	5.49	5.10	4.49
Not now, used to smoke daily	11.26	10.30	10.31	12.82
Not now, used to smoke occasionally	5.08	4.71	4.23	4.45
Never smoked	48.92	51.29	53.01	50.41

Source: Authors´ elaboration from ECHP.

In order to establish the main socio-demographic characteristics of smokers, we have classified them into seven groups of variables: personal characteristics, education level, marital status, income, occupational status and other variables related to individuals' health, household characteristics and social relationships. Table 4 shows explanatory variables used in estimations and their corresponding definitions.

Table 4: Variables Definitions

Variable Name	**Variable Definition**
Personal Characteristics	
Gender (MALE)	1 if male, 0 otherwise
Age (AGE)	Age in years at 31st December of current wave
Education Level	
Illiterate (LOWEDUC)	1 if highest academic qualification is less than secondary level (ISCED 0-2), 0 otherwise
Marital status	
Married (NVRMAR)	1 if married, 0 otherwise
Separated (SEPARATED)	1 if separated, 0 otherwise

Divorced (DIVORCED)	1 if divorced, 0 otherwise
Widow (WIDOW)	1 if widowed, 0 otherwise
Income	
Net Income (LINCOMEOCDMO)	Logarithm of equivalised annual household net income (OECD modified scale)
Occupational Status	
Status in employment (UNEMPLOYED)	1 if unemployed, 0 otherwise
Health Status	
Self-assessed health (GOODSAH)	1 if individual's self-assessed health is good of very good, 0 otherwise
Household	
Household size (HHSIZE)	Number of people in household including respondent
Social Relationships	
Personal relationships (SOCIALCL)	1 if member of a club or organisation, 0 otherwise

Source: Authors´ elaboration from ECHP.

Firstly, as personal characteristics we have included two variables: individual's age (in years) and gender (building a dummy variable which takes value of 1 if individual is male and 0 otherwise).

The second group of variables are referred to the maximum level of education completed. In the ECHP, education is classified into three categories based on ISCED classification: less than secondary level (ISCED 0-2), second stage of secondary level (ISCED 3) and third level (ISCED 5-7). Thus, a dummy variable has been included: less than secondary level (LOWEDUC). In fact, the prevalence and amount of tobacco smoking is concentrated among men and women with lower education, lower income and lower occupational class (Cavelaars et al. 2000, Osler et al. 1998). In this way, poor socio-economic conditions in youth and adolescence influence uptake through a range of mechanisms, including decreased refusal skills and increased psychosocial stress (Kunst et al. 2004).

Thirdly, representing marital status, we have considered four variables (married, separated, divorced and widow) with never married as the reference category.

On the other hand, we are concerned with the influence of income on smoking decision. Our income variable is equivalised annual net household income (LINCOMEOCDMO) adjusted using OECD modified scale to take into account household size and composition. In this sense, we have used household information rendering the component family by using equivalence scales. The modified OECD scale gives a weight of 1 to the first adult, 0.5 to other persons aged 14 or over and 0.3 to each child aged less than 14. For each person, the "equivalised total net income" is calculated as its household total net income divided by equivalised household size. In this case, we use the logarithm of household´s income (OECD modified scale).

Other variables included in the analysis related to occupational status are status in employment. We have considered a dummy variable that takes the value one if the individual is unemployed and zero otherwise (UNEMPLOYED).

Also, we have considered other variables related to health status. The variable GOODSAH indicates whether or not the individual's self-assessed health is good or very good.

Finally, we have considered number of people in household including respondents (Household size-HHSIZE). Also, we have included variables related to social relationships, and another dummy variable has been built in order to take into account whether an individual is a member of a club or organisation (SOCIAL) or not.

Table 5 reports the results of the estimation including two types of explanatory variables. The first type can be treated as though they were continuous variables (individual's age measured in years and household income) and other explanatory variables are binary or dummy variables. These take the value 1 if the individual has a particular characteristic. In this way, the marginal effects make us aware of the impact of a small change in the variable on the probability of participation. Thus, we can study the impact of age on the probability of smoking. On the other hand, for the dummy variables, it does not make sense to think in terms of small changes (an individual either has a characteristic or does not). So, we will consider the average effects, that is, for example, the difference in the probability of

being a smoker if someone is unemployed compared to someone who is employed.

Table 5: Probit Estimates including average and marginal effects. Dependent variable: Dummy variable which takes value one if individual smokes daily or occasionally.

	Explanatory variables	**Coef. (Std.Err.)**	**dF/dx (Std.Err.)**
Personal Characteristics	MALE	0.5382 (0.0263)	0.1833 (0.0088)
	AGE	-0.0206 (0.0011)	-0.0070 (0.0003)
Education level	ANALFA	0.0322 (0.0295)	0.0110 (0.0100)
Marital Status	MARRIED	0.3321 (0.0341)	0.1112 (0.0111)
	SEPARATED	0.9528 (0.1066)	0.3646 (0.0395)
	DIVORCED	0.8567 (0.1284)	0.3287 (0.0490)
	WIDOW	0.0511 (0.0768)	0.0177 (0.0268)
Income	LINCOMEOCDMO	0.0071 (0.0188)	0.0024 (0.0064)
Occupational Status	UNEMPLOYED	0.2266 (0.0508)	0.0813 (0.0190)
Health status	GOODSAH	0.1505 (0.0338)	0.0505 (0.0111)
Household	HHSIZE	-0.0238 (0.0095)	-0.0081 (0.0032)
Social Relationships	SOCIAL	-0.0397 (0.0301)	-0.0135 (0.0101)
Number of obs	11452		
Log likelihood	-6431.3136		

SOURCE: Authors´ elaboration from ECHP (2000 and 2001).

The sign of the coefficients inform us about the qualitative effect of the explanatory variables. In this way, if the sign of the coefficient on UNEMPLOYED is positive, this means that an individual who is currently "Unemployed" is more likely to be smoker relative to the reference individual who is employed. Estimates show that most of the coefficients are significant and have the expected signs. Individuals who are more likely to be smoker are unemployed young men with low level of education.

Conclusions

In Spain from 1st January 2006, new laws ban smoking in public places including food shops and all workplaces. On the other hand, smoking areas should be set aside in hotels, airports, stations or ports. One of the objectives of the law is to protect the rights of non-smokers. However, it supposes a cultural and lifestyle change in Spain.

This study analyses the relationship between health expenditure, life expectancy, gross domestic product and tobacco mortality in Spain. We can confirm that these variables are not integrated with the same order so the causality effect, from a statistical point of view, is not so clear. Furthermore, using the information contained in the European Community Household Panel, we can observe that men with lower educational background and unemployed are more likely to smoke. On the other hand, it should be taken into account that in many cases women are "social smokers", that is, smoking has become a sign of Spanish women's emancipation. In fact, men and women from lower socio-economic groups have a higher risk of initiating smoking and become addicted during adolescence. These individuals are important target groups for anti-smoking policies. Therefore, inequalities in both smoking initiation and cessation produce large socio-economic differences in the life time exposure to smoking (Kunst et al. 2004). Even socio-economic differences in mortality can be explained by inequalities in smoking (Mackenbach et al. 2004).

Thus, Spain's anti-smoking battle is very important in order to reduce inequalities through tobacco control policies. Increases in tobacco taxes are one of the most effective tool for reduction in tobacco consumption, especially among young people. This measure must be accompanied by advertising bans, anti-tobacco mass media campaigns,

consumer information, etc. Obviously, rising tobacco prices could have larger effects among lower socio-economic groups than among higher ones.

References

[1]. Bask, M., Melkersson M. (2003). *Should one use smokeless tobacco in smoking cessation programs? A Rational Addiction Approach*. European Journal of Health Economics 4, 263-270.

[2]. Cavelaars, A., Kunst, A., Geurts, J.J.M., Crialesi, R., Grötvedt, L., Helmert, U., Lahelma, E., Lundberg, O., Matheson, J., Mielck, A., Rasmussen, NKr., Regidor, E., Rosário-Giraldes, M., Spuhler, Th., Mackenbach, J.P. (2000) *Educational differences in smoking: International comparison*. British medical Journal 320, 1102-107.

[3]. Davidson R, MacKinnon J G (2006). *Bootstrap methods in econometrics. In Palgrave. Handbook of Econometrics*, Volume 1, T.C. Mills and K. Patterson (eds.). Palgrave MacMillan, Basingstoke.

[4]. Dickey DA, Fuller WA (1979). *Distribution of the Estimators for Autoregressive Time Series with a Unit Root*. J. Am. Stat. Assoc. (74): 427 - 731.

[5]. European Network for Smoking Prevention project report, (2004). *Tackling socio-economic inequalities in the European Union.* www.ensp.org/projects.cfm#200022003Projects.

[6]. Graham, H. (1996). *Smoking prevalence among women in the European Community 1950-1990.* Social Science and Medicine 43, 243-254.

[7]. Granger CWJ (1981). *"Some Properties of Time Series and Their Use in Econometric Model Specification*. J. Health. Econ. (16): 121 - 130.

[8]. Greene, W.H. (2003). *Econometric Analysis. Prentice Hall.*

[9]. Jacobzone, S. (2000), "*Pharmaceutical Policies in OECD Countries: Reconciling Social and Industrial Goals*", OECD Labour Market and Social Policy Occasional Papers, No. 40, OECD Publishing. http://dx.doi.org/10.1787/323807375536.

[10].Johansen S (1991). *Estimation and hypothesis testing of cointegration vectors*. Econometrica, (59): 1551 - 1580.

[11]. Jones, A. (2001). *Applied Econometrics for Health Economists-A practical guide*. Office of Health Economics, London.

[12]. Kunst, A., Giskes, K., Mackenbach, J. (2004). *Socio-Economic Inequalities in Smoking in the European Union. Applying an Equity Lens to Tobacco Control Policies. EU Network on Interventions to Reduce Socio-Economic Inequalities in Health.*

[13]. Mackenbach, J.P., Huisman, M., Andersen, O. (2004). *Inequalities in lung cancer mortality by the educational level in 10 European populations*. European Journal of Cancer 40(1), 126-135.

[14]. Maddala GS (1992). *Introduction to Econometrics*. MacMillan Publishing Company, New York.

[15]. Osler, M., Prescott, E., Gottschau, A., Bjerg, A., Hein, H., Sjol, A. and Schnohr, P. (1998) *Trends in Smoking Prevalence in Danish Adults 1964-1994. The Influence of Gender, Age and Education*. Scandinavian Journal of Social Medicine 26, 293-298.

[16]. Regidor, E., Gutiérrez-Fisac, M.D., Calle, M.E., Navarro, P., Domínguez, V. (2001). *Trends in cigarette smoking in Spain by social class*. Preventive Medicine 33, 241-248.

[17]. Ross, H. (2004). *The Economics of Tobacco and Tobacco Control in the European Union in Tobacco or Health in the European Union: Past, Present and Future*, Chapter 2, 69-98, European Commission.

[18]. Rovira, J., W. Kip Viscusi & Antoñanzas; F., Costa-i-Font, J., Hart, W., Carvalho, I. (2000). "*Smoking risks in Spain: part II - perceptions of environmental tobacco smoke externalities*," LSE Research Online Documents on Economics 13318, London School of Economics and Political Science, LSE Library.

[19]. Phillips PCB, Perron P (1988). *Testing for a Unit Root in a Time Series Regression*. Biometrika, (75): 335 - 346.

[20]. Sato, M., Ohkusa, Y. (2003). *The relationship between smoking initiation and time discount factor, risk aversion and information.* Applied Economics Letters 10, 287-289.

[21]. Suranovic, S.M., Goldfarb, R.S., Leonard, T.C. (1999). *An economic theory of cigarette addiction*. Journal of Health Economics 18, 1-29.

[22]. Waldron, I. (1991). *Patterns and causes of gender differences in smoking.* Social Science and Medicine 32, 989-1005.

[23]. World Health Organization (WHO) (2004). Part 1: *Tobacco as a risk factor: Health, social and economic costs in Building Blocks on Tobacco Control: A Handbook*. WHO Cataloguing- in – Publication Data, France.

[24]. World Health Organization (WHO) (2012). *Mortality Attributable to Tobacco, Global Report*.