Smoking, health related quality of life and economic evaluation

Authors: Ángel López-Nicolás\textsuperscript{a*}, Marta Trapero-Bertran\textsuperscript{bc}, Celia Muñoz\textsuperscript{c}

\textsuperscript{a} Departamento de Economía, Universidad Politécnica de Cartagena (UPCT)
\textsuperscript{b} Universitat Internacional de Catalunya (UIC). Faculty of Economic and Social Sciences
\textsuperscript{c} Centre de Recerca en Economia i Salut (CRES), Universitat Pompeu Fabra (UPF)

Ángel López-Nicolás
Universidad Politécnica de Cartagena
Faculty of Business Science
Economics Department
C/ Real, nº 3
30201 Cartagena (Murcia)
Spain

Marta Trapero-Bertran and Celia Muñoz
Universitat Internacional de Catalunya
Faculty of Economic and Social Sciences
c/Immaculada 22
08017 Barcelona (Catalonia)
Spain

Corresponding author:
Ángel López-Nicolás (angel.lopez@upct.es);
Tel. 968 325484 / 968 325491
Fax. 968 325781.
Abstract

Background and aims

The economic evaluation of tobacco control policies requires the adoption of assumptions about the impact of changes in smoking status on health related quality of life (HRQOL). Estimates for such impacts are necessary for different populations. This paper aims to test whether smoking status has an independent effect on HRQOL over and above the effect derived from the increased likelihood of suffering a tobacco related disease and to calculate utility values for the Spanish population.

Methods

Using data from the Spanish Encuesta Nacional de Salud of 2011-12, we estimate statistical models for HRQOL as measured by the EQ5D5L instrument as a function of smoking status. We include a comprehensive set of controls for biological clinical, lifestyle and socioeconomic characteristics.

Results

Smoking status has an independent, statistically significant effect on HRQOL. However, the size of the effect is small. The typical smoking related disease, such as lung cancer, is associated to a reduction in HRQOL about 5 times larger than the difference between current smokers and never smokers.

Conclusion

A realistic representation of the effects of smoking on HRQOL in economic evaluation should shy away from attributing large independent changes to quitting smoking or avoiding starting to smoke, since such changes are small once clinical conditions are controlled for. On the other hand, it is necessary to expand the set of classical smoking related diseases used in economic evaluation with other diseases for which new evidence showing a causal link to smoking exists.

Keywords (max 6): smoking; quality of life; economic evaluation; EQ5D5L

JEL: I12 Health Behaviour; D61 Allocative efficiency – Cost-Benefit Analysis; (https://www.aeaweb.org/econlit/jelCodes.php?view=jel#I)
**Acknowledgments:**

The authors are very grateful to Dr. Toni Mora for his invaluable help and support to finish this paper. This work was partially funded by the European Union’s FP7 program (The EQUIPT Project; grant agreement 602270) and partially by the Spanish Ministry of Economy and Competitiveness (MINECO) under the programme "Programa Estatal de Investigación, Desarrollo e Innovación Orientada a los Retos de la Sociedad, Plan Estatal de Investigación Científica Técnica y de Innovación 2013-2016”. Grant ECO2013-48217-C2-1 (http://invesfeps.ulpgc.es/en). The funders had no influence in the conduction of this study or the drafting of this manuscript.
Smoking, health related quality of life and economic evaluation

Introduction

That tobacco causes disease is a long established fact. Globally 12% of all deaths among adults aged 30 years and over were attributed to smoking [1]. Tobacco kills around 6 million people each year. More than 5 millions of those deaths are the result of direct tobacco use while more than 600,000 are the result of non-smokers being exposed to second-hand smoke [2]. The list of health conditions for which there exists scientific evidence showing a causal effect is likely to continue to grow, with the latest report of the Surgeon General [3] adding diabetes mellitus, rheumatoid arthritis, colorectal cancer as well as general inflammation and impairment of the immune system to the “classical” group of smoking related ailments such as lung cancer, chronic obstructive pulmonary disease (COPD), myocardial infarction, coronary disease, or stroke. Economic evaluations of tobacco control policies typically account for the loss of quality of life associated to suffering these diseases by means of health related quality of life (HRQOL) indices that permit the calculation of some measure of quality adjusted life years such as the well-known QALYs [4,5]. A relevant research question, with important implications for policy, is whether smoking affects HRQOL over and above its effect on the likelihood of contracting disease. As Vogl et al. [6] have argued, smoking may induce changes in utility in individuals who are otherwise equal to non-smokers in terms of biological, clinical and social characteristics. Such changes need to be duly accounted for in cost-effectiveness, cost-utility and general return on investment metrics for tobacco control policies.

The main aim of this paper is to contrast whether the smoking status of the general Spanish population is associated to systematic variations in HRQOL as measured by the EQ5D5L instrument [7,8] once biological and clinical conditions are controlled for. The focus on this HRQOL instrument responds to its widespread use in economic evaluations of tobacco control policies [9-12]. While there are previous studies using Spanish data [13-16], these use the SF-36 instrument on a small sample or do not control for comorbidities. Up to our knowledge, there is only one cost-utility analysis for Spain [16]. Moreover, while these and other international studies document differences in HRQOL by smoking status, they do not control exhaustively for clinical conditions potentially correlated with tobacco consumption so it is difficult to attribute an independent effect
on HRQOL to smoking. In contrast, our use of the Spanish National Health Survey, containing a sample of around 20000 individuals from the general population reporting a wide array of clinical conditions, offers the possibility of contrasting for the existence of such independent effects with some degree of confidence.

Methods

Our data source is the latest National Health Survey release, that of year 2011-12 [17]. The ENS11-12 is representative sample of the non-institutionalized Spanish population containing information on lifestyles, health and socioeconomic characteristics of individuals with separate adults (16+) and children samples. The analysis in this paper is based on the adult sample. The ENS2011-12 is the first ENS that contains information on the EQ5D5L health instrument [7,8].

In common with other populations, the distribution of the EQ5D5L has a substantive concentration of cases at the maximum tariff (71.4% for men and 54.8% for women, see Figure 1). Other noteworthy characteristics of the EQ5D5L distribution in the ENS2011-12 are i) the presence of small percentage of cases with tariffs below zero (1.9% for men and 4.7% for women), which in the EQOL literature are traditionally interpreted as “worse than death” health states [18] and ii) a high degree of right skewness in the distribution of uncensored tariffs.

Our statistical analysis hinges on the specification of models that aim to explain the variation in the EQ5D5L tariff as a function of biological and clinical characteristic and lifestyles. These models need to account for the censored and skewed nature of the dependent variable. Similarly, the remarkable differences between males and females mentioned above call for a separate analysis for both genders. Among the various statistical alternatives suggested in the literature we opt for the Two Part Model (TPM) [19]. The first part of the model estimates the probability of reporting no health problems by means of a Probit Model. The second part explains the expectation of the health utility score (tariff) given that some health problem has been reported by means of a Generalized Linear Model (GLM) with a logarithmic link and gamma disturbances. The TPM has been shown to produce good results in terms of predictive power in comparison with other models in the context of the EQOL5D [19]. Also, the TPM is readily interpretable. As
mentioned above the first part serves to predict the probability of reporting no health problems (which below will be referred to as \( P(\text{no health problems reported}) \)) and the second part serves to predict the expected value for the tariff conditional on reporting a health problem, denoted as \( E(\text{tariff| some health problem reported}) \) below.

With regard to the explanatory variables, we use 6 different specifications or models. Our baseline specification, Model 0, contains indicators for smoking status distinguishing between current smokers, former smokers and never smokers, a quadratic polynomial in age and controls for marital status, levels of alcohol consumption, physical activity, body mass index and an indicator for exposure to second hand smoke. Models 1 to 4 add alternative sets of explanatory variables to the baseline specification. Namely, Model 1 includes indicators for a medical diagnosis of heart infarction, malignant tumor, coronary obstructive pulmonary disease (COPD), stroke, other heart diseases and asthma (five of the classical tobacco related diseases). Model 2 includes indicators for mental disorders, such as depression, anxiety or other mental problems. Model 3 includes pain conditions such as migraine, back pain, arthritis and recent injuries. Model 4 contains other medical diagnosis such as hypertension, varicous veins, allergy, diabetes, stomach ulcer, urinary incontinence, high cholesterol, cataracts, skin problems, constipation, liver cirrhosis, hemorrhoids, osteoporosis, thyroid problems, menopausal problems (for women), prostate problems (for men). Finally, the full specification, Model 5, adds all the indicators used in Models 1, 2, 3 and 4 to the baseline specification.

The rationale behind these specifications was the necessity to test whether any systematic association of smoking status with HRQOL is robust to the inclusion of different sets of clinical conditions. In the case of specifications 2 and 3, which add, respectively, mental problems and pain, the test is particularly demanding, in the sense that two of the EQ5D5L domains are precisely mental problems and pain. Of course, in the context of a cross section of non-experimental observational data we cannot rule out that such effects, if they exist, are due to correlated unobservables. In order to explore this possibility we carry out a robustness check consisting in expanding specification 5 with controls for social class and degree of perceived social capital.

From these two components it is possible to retrieve the predictions for the unconditional expectation of the tariff, simply as:
\[ E(\text{tariff}) = P(\text{no health problems reported}) \times \text{value of maximum tariff} + (1 - P(\text{no health problems reported})) \times E(\text{tariff} \mid \text{some health problem reported}). \]

These unconditional expectations, and their conditional (on reporting some health problem) counterparts, i.e. \( E(\text{tariff} \mid \text{some health problem reported}) \), may be used to produce estimates for the EQOL5D based HRQOL tariff of prototypical profiles of individuals by gender, smoking status and age to use in cost-utility analysis of tobacco policies.

Results

Figure 2 (Graph a) shows the relationship between smoking status and the fraction of individuals reporting no health problems, with the mean age for each subgroup displayed at the top of the corresponding bar. In the case of men, the highest fraction is found amongst never smokers, followed by current smokers. Former smokers are more likely to report some health problem. These differences, which do not control by any other covariates, are partly explained by differences in age across smoking status. Male former smokers are more than 10 years older on average than current and never smokers. In the case of women never smokers are more likely to report some health problem and current smokers are the least likely group to report problems but, again, these differences seem to be driven by differences in average age. In contrast to men, female never smokers are older than either current or former smokers.

Graph b in Figure 2 shows the corresponding breakdown for the average EQ5D5L tariff for those that report some health problem. Not surprisingly, these individuals are older on average for all smoking states and genders. But a similar pattern to that shown in Graph a of Figure 2 emerges whereby male former smokers and female never smokers, the oldest age groups within their respective genders, report the worst levels of health.

Table 1 presents the estimates for the marginal effects of smoking status on HRQOL within the TPM for the models described above, along with the Akaike Information Criterion (AIC) measure of goodness of fit. The top panel corresponds to Part 1 of the TPM, that is, the probability of reporting no health problems in any of the EQOL5D domains, while the bottom panel corresponds to the second part of the TPM, i.e. the model for expected value of the tariff conditional on reporting some health problem. The omitted category within the smoking status set of dummy variables is “never smoker”.
For the first part of the TPM, note that the best specification in terms of the AIC statistic is the one containing the full set of explanatory variables (Model 5), both for males and females. Among Models 1-4, which add alternative sets of covariates to the baseline specification in Model 0, the one including pain conditions (Model 4) results in the best improvement in goodness of fit with respect to the baseline specification, followed by the model including mental diseases (Model 2). This is not surprising since mental disorders and pain are two of the dimensions along which the EQOL5D5L is measured. The inclusion of tobacco related diseases (Model 1) improves the AIC with respect to the baseline specification by smaller margins.

The marginal effect of current smoking on the probability of reporting some health problem among both males and females ranges between 4% and 2%, this latter estimate corresponding to the best performing model (Model 5), which in the case of women verges on statistical insignificance (p value=0.109). As for the marginal effect of former smoking, it ranges between 5% in the baseline specification and statistical insignificance for both men and women in Model 5.

For the second part of the TPM, Model 5, in similarity to the rest of specifications, results in no clear improvements in the AIC with respect to the baseline specification. In the case of males, the marginal effects of the smoking status variables do not exert a significant effect on the expected level of the tariff (conditional on having reported a health problem). For females we find a significant but small (about 0.02 EQ5D5D tariff points) marginal effect for being a former smoker.

These results are robust to the inclusion of controls for social class and degree of perceived social support.

Table 2 presents estimates for the expected EQ5D tariff for a set of representative profiles broken down by age, gender and smoking status. These estimates are defined as the unconditional expectation of the tariff over the relevant population group, and they have been calculated with the two parts of Model 5. Note that, within age and gender categories, there are no stark differences in the expected EQOL5D5L tariff by smoking status.

On the other hand, Table 3 presents estimates for the change in the tariff associated to suffering a tobacco related disease. They are defined as the difference between the
unconditional expectation of the HRQOL score over the population of individuals who do not suffer any of such diseases minus the expectation of the HRQOL conditioned on suffering the corresponding disease and reporting health problems for the same population. Note that for some diseases this change is very substantial. For instance, the drop in the tariff reaches about 0.35 tariff points in the case of stroke.

Discussion

The conjunction of results presented above suggests a series of stylized facts about the relationship between smoking and HRQOL as measured by the EQ5D5L. First, even the most comprehensive specifications in terms of clinical, biological and lifestyle conditions detect an independent effect of smoking on HRQOL in comparison to otherwise equal never smokers. This effect operates through a larger probability of reporting some health problem, but not through current smokers reporting a lower tariff than otherwise equal never smokers who also report health problems along any of the EQ5D5L dimensions.

We find that being a former smoker also seems to affect the probability of reporting health problems, but its effect is not statistically significant once the full set of available reported clinical diagnoses is included. This suggests that the former smoker status is a proxy for clinical diagnoses. In the case of women, though, we find that being a former smoker has a small and significant negative effect on the expected EQ5D5L tariff among those who report a health problem. This gender effect is probably a result of the differences in the evolution of the smoking epidemic in Spain, where for male former smokers the average period since quitting is longer than for female former smokers (for instance, the proportion of male former smokers who quit more than 10 years before the date of the survey is 56% while the corresponding figure for females is 42.4%).

Nonetheless, the effects of smoking on HRQOL are very small in magnitude once clinical conditions are comprehensively controlled for. For instance, currently smoking women in the 45-54 age band are expected to have a EQ5D5L tariff of 0.89 compared to a tariff of 0.92 for women in the same age band who have never smoked or 0.91 for former smokers.

In contrast, the substantive damaging effect of smoking operates through the reduction in HRQOL associated to suffering a smoking related disease. For instance, having a stroke
reduces the EQ5D5L by a margin more than ten times larger than the difference between current and never smokers mentioned in point 3 above. For those that suffer a heart infarction, other heart diseases, COPD or a tumor the margin is about 5 times larger, and for asthma the difference is about 3 times larger.

There are two stark implications of these results for research on the cost-effectiveness, the cost-utility and the return on investment in general of tobacco control policies. Firstly, attributing substantive HRQOL gains to quitting smoking as well as accounting for the concomitant HRQOL gain derived from a smaller likelihood of contracting tobacco related diseases might lead to an overestimation of the benefits of tobacco control policies. Second, but not least, there will be an underestimation of such benefits to the extent that policy analysts omit diseases for which tobacco might have been a causal factor. For instance, a diagnosis of either arthritis or diabetes, two diseases causally associated to smoking according to the latest report from the General Surgeon, but nonetheless typically omitted in economic evaluation of tobacco policy, are associated to a reduction of about 0.15 in HRQOL as measured by the EQOL5D tariff. This effect is about 5 times larger than the difference between smoking currently and not having smoker ever for women in the 45-54 age band. New economic evaluation research in the area of tobacco should consider the inclusion of such diseases.
References


Table 1  Marginal effect estimates for smoking status in Two Part Models. Omitted category: never smoker.

<table>
<thead>
<tr>
<th>Part 1: Prob (No health problems reported)</th>
<th>Probit Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEN (N=9619)</td>
</tr>
<tr>
<td></td>
<td>Model 0</td>
</tr>
<tr>
<td>CURRENT</td>
<td>-0.0379***</td>
</tr>
<tr>
<td>FORMER</td>
<td>-0.0531***</td>
</tr>
<tr>
<td>AIC</td>
<td>1.044</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part 2: E(Tariff</th>
<th>Some health problem reported)</th>
<th>Generalized linear model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEN (N=9619)</td>
<td>WOMEN (N=11337)</td>
</tr>
<tr>
<td></td>
<td>Model 0</td>
<td>Model 1</td>
</tr>
<tr>
<td>CURRENT</td>
<td>0.0010</td>
<td>0.0027</td>
</tr>
<tr>
<td>FORMER</td>
<td>-0.0088</td>
<td>-0.0034</td>
</tr>
<tr>
<td>AIC</td>
<td>1.529</td>
<td>1.531</td>
</tr>
</tbody>
</table>

* denote p values between 0.05 -0.1; **denote p values between 0.01-0.05; ***denote p values below 0.01
Table 2  Estimates for the unconditional expectation of the EQOL5D5L tariff by age, gender and smoking status, with bootstrapped standard errors

<table>
<thead>
<tr>
<th>Status/age band</th>
<th>Men</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Status/age band</td>
<td>15-24</td>
<td>25-34</td>
<td>35-44</td>
<td>45-54</td>
<td>55-64</td>
<td>65-74</td>
<td>75-84</td>
<td></td>
</tr>
<tr>
<td>Current smokers</td>
<td>0.951</td>
<td>0.953</td>
<td>0.949</td>
<td>0.931</td>
<td>0.921</td>
<td>0.899</td>
<td>0.834</td>
<td></td>
</tr>
<tr>
<td>Former smokers</td>
<td>0.971</td>
<td>0.962</td>
<td>0.959</td>
<td>0.941</td>
<td>0.915</td>
<td>0.896</td>
<td>0.816</td>
<td></td>
</tr>
<tr>
<td>Never smokers</td>
<td>0.964</td>
<td>0.967</td>
<td>0.962</td>
<td>0.948</td>
<td>0.922</td>
<td>0.916</td>
<td>0.83</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Status/age band</th>
<th>Women</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Status/age band</td>
<td>15-24</td>
<td>25-34</td>
<td>35-44</td>
<td>45-54</td>
<td>55-64</td>
<td>65-74</td>
<td>75-84</td>
<td></td>
</tr>
<tr>
<td>Current smokers</td>
<td>0.947</td>
<td>0.938</td>
<td>0.929</td>
<td>0.896</td>
<td>0.873</td>
<td>0.822</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Former smokers</td>
<td>0.965</td>
<td>0.953</td>
<td>0.948</td>
<td>0.916</td>
<td>0.85</td>
<td>0.818</td>
<td>0.662</td>
<td></td>
</tr>
<tr>
<td>Never smokers</td>
<td>0.965</td>
<td>0.963</td>
<td>0.948</td>
<td>0.92</td>
<td>0.855</td>
<td>0.791</td>
<td>0.672</td>
<td></td>
</tr>
</tbody>
</table>
Table 3  Estimates for the expectation of the EQOL5D5L tariff conditional on reporting health problems and reporting a diagnosis of a selection of diseases by gender, with bootstrapped standard errors.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>COPD</td>
<td>-0.159</td>
<td>-0.144</td>
</tr>
<tr>
<td></td>
<td>0.022</td>
<td>0.021</td>
</tr>
<tr>
<td>Stroke</td>
<td>-0.356</td>
<td>-0.367</td>
</tr>
<tr>
<td></td>
<td>0.047</td>
<td>0.049</td>
</tr>
<tr>
<td>Heart infarction</td>
<td>-0.176</td>
<td>-0.152</td>
</tr>
<tr>
<td></td>
<td>0.029</td>
<td>0.04</td>
</tr>
<tr>
<td>Other heart disease</td>
<td>-0.159</td>
<td>-0.165</td>
</tr>
<tr>
<td></td>
<td>0.02</td>
<td>0.019</td>
</tr>
<tr>
<td>Tumor</td>
<td>-0.194</td>
<td>-0.139</td>
</tr>
<tr>
<td></td>
<td>0.032</td>
<td>0.022</td>
</tr>
<tr>
<td>Arthritis</td>
<td>-0.201</td>
<td>-0.192</td>
</tr>
<tr>
<td></td>
<td>0.015</td>
<td>0.02</td>
</tr>
<tr>
<td>Diabetes</td>
<td>-0.192</td>
<td>-0.141</td>
</tr>
<tr>
<td></td>
<td>0.02</td>
<td>0.012</td>
</tr>
</tbody>
</table>
Fig. 1 EQOL-5D-5L by gender

Graph a
EQOL5D5L values for women

Graph b
EQOL5D5L values for men
Fig. 2  Smoking status and EQOL-5D-5L: Proportion of values censored at 1 and mean score for uncensored range

Graph a
Proportion of values censored at 1 and the mean age in group

Graph b
Mean score uncensored range and mean group age

**Average age for group displayed on top of the corresponding bar**