UNIVERSITÀ DEGLI STUDI DI VERONA

WORKING PAPER SERIES

Addiction, Social Interactions and Gender Differences in Cigarette Consumption

by

David Aristei and Luca Pieroni

June 2007 - #39

DIPARTIMENTO DI SCIENZE ECONOMICHE

ADDICTION, SOCIAL INTERACTIONS AND GENDER DIFFERENCES IN CIGARETTE CONSUMPTION

David ARISTEI[1] – Luca PIERONI[2]

Abstract

This paper addresses the impact of addiction and social interactions on cigarette demand, controlling for demographic and socioeconomic factors. A Box-Cox double-hurdle model for the simultaneous decisions of how much to smoke and whether to quit smoking is estimated on individual data from the 2000 Italian "Health Status and Use of Health Services" survey. The model incorporates the fixed costs of quitting and allows for the analysis of the effects of addiction and within-household interactions on smoking participation and cigarette consumption. Estimation results show that the duration of the smoking habit, used as measure of addiction, significantly increases the level of cigarette consumption and lowers the probability of quitting. Within-household social interactions affect individual's attitude toward smoking. Participation decision is significantly influenced by the presence of other smokers and individual cigarette consumption increases as the consumption of the peer-group grows. Finally, gender differences are formally tested to verify whether male and female sub-samples can be pooled or should be separately analyzed. The hypothesis of equal consumption parameters is clearly rejected, suggesting the opportunity of distinguishing the consumption patterns of men and women.

J.E.L. classification: C24, D12, J16

Keywords: cigarette consumption, social interactions, gender effects, double-hurdle models.

[1] *David Aristei: Department of Economics, University of Verona and Department of Economics, Finance and Statistics, University of Perugia, e-mail: daristei@economia.univr.it.*

[2] *Luca Pieroni: Department of Economics, Finance and Statistics, University of Perugia, via Pascoli 20, 06123 Perugia (Italy); Tel. +390755855280, Fax +390755855299, e-mail: lpieroni@unipg.it.*

1 Introduction

The demand for addictive goods is characterized by complex determinants affecting the decisions of participation and consumption. The purpose of this paper is to disentangle the impact of addiction, demographics and social interactions on cigarette demand, in order to explain the level of cigarette consumption and the decision to quit smoking.

Among the main reasons for studying cigarette consumption, an explicit investigation of smoker's addictive behaviour is useful to correctly model health policies. The framework proposed in this paper follows the works of Suranovic *et al.* (1999) and Jones (1999) in modelling quitting costs, according which persistent patterns of individual smoking behaviour generate significant "fixed costs" in the reduction (or elimination) of cigarette consumption.

On the other hand, demographic and socio-economic factors are proved to separately influence both participation and cigarette consumption decisions and play an active role in modifying individual preferences (Jones, 1994; Jimenez-Martin *et al.*, 1998; Yen, 2005a). A large part of the recent literature finds that participation and consumption decisions can also be conditioned by social interactions (Brock and Durlauf, 2001; Powell *et al.*, 2005; Krauth, 2005). In analyzing smoking behaviour, the underlying assumption is that the utility that an individual receives from smoking depends on the actions of the other individuals within the person's reference group (Krauth, 2005).

Finally, the role of demographic components in modifying individual preferences are investigated by explicitly considering the hypothesis of gender differences in cigarette consumption. The existence of gender-differentiated behaviours in consumption has been empirically proved to be relevant, suggesting specific health policies for men and women (Yen, 2005; Aristei and Pieroni, 2007).

Our empirical strategy aims to address the simultaneous decisions of how much to smoke and whether to quit smoking by using of a flexible Box-Cox double-hurdle model (Yen, 1993; Yen and Jones, 2000). The specification proposed generalizes the structure of several nested univariate and bivariate models and incorporates the effects of addictive behaviour and within-household interactions on smoking participation and cigarette consumption, while controlling for individual heterogeneity. In this way we can encompass, besides the standard double-hurdle model, a wide range of specifications, that differ for the distributional assumptions on the error terms, and test the best model to rationalize the data. Moreover, differently from other studies carried out for Italy (Aristei and Pieroni, 2007), we use individual rather than household's data. The use of individual data from the 2000 Italian "Health Status and Use of Health Services" survey allows us to address the problem of measurement error in estimating separately the determinants of participation and consumption.

The paper is organized as follows. The next section outlines the theoretical framework upon which the empirical models are based. In Section 3 we discuss econometric methods. Special emphasis is placed on the extension of the Box-Cox specification to account for within-household social interactions. In Section 4.1 the data used in the empirical analysis, taken from the 2000 Italian Health Status and Use of Health Services survey, are discussed and factors influencing participation and consumption equations are examined. Sections 4.2 and 4.3 discusses specification and estimation results are presented, with specific attention devoted to the nested strategy used to obtain the best model and to analysis the estimated parameters and elasticities. In Section 4.4, we deepen the analysis of gender-differentiated consumption behaviours and present the results obtained for the men and women sub-samples. Finally, Section 5 concludes the paper with a summary of the work and a discussion on health policy implications.

2 Theoretical framework

2.1 A model for smoking and quitting decisions

In this section, a model for the analysis of the simultaneous decisions of how many cigarettes to smoke and whether to quit smoking, based on the works of Jones (1994) and Yen and Jones (1996), is derived. The framework of analysis rests on the trade-off between the expected benefits of quitting and the fixed costs of quitting associated with the effects of nicotine dependence and withdrawal. Even if we adopt a static framework, the specification is intended to model the expected future benefits of quitting and could be therefore interpreted in terms of intertemporal models of addiction. Previous studies of addiction based on individual data (Chaloupka, 1991; Labeaga, 1993, 1999) have focused on the effect of addiction on current cigarette consumption levels. In this study, since the dataset used provides information on individuals' smoking history, irrespective of whether they are current or ex smokers, we are able to separately analyze the influence of addiction on both quitting and consumption decisions.

Referring to Becker and Murphy's (1988) rational addiction model, the development of nicotine dependence can be characterized in terms of tolerance, reinforcement and withdrawal effects. In particular, the rational addiction approach suggests that addiction arises whenever consumption of a good displays adjacent complementarity, so that an individual is addicted to a good only when past consumption raises the marginal utility of present consumption (reinforcement effect). However, an implication of adjacent complementarity is a negative utility effect that occurs whenever an individual with a past consumption stock attempts to reduce present consumption (withdrawal effect). In the model developed here, we incorporates this negative utility effect explicitly and interpret it as a quitting cost, as in Suranovic *et al.* (1999).

Two features of withdrawal effects should be remarked. Firstly, the effects are asymmetric and only occur when smokers try to reduce consumption or quit. Secondly, especially for heavy smokers, consumption not only provides satisfaction but it is also necessary to avoid the consequences of withdrawal. The influence of the withdrawal effects can be modelled as an adjustment cost, which represents the discomforts that arise whenever cigarette consumption is reduced or eliminated. In line with the framework proposed by Jones (1994) and Yen and Jones (1996), we assume that these adjustment costs (A_i) depend on the amount, duration and pattern of past consumption as well as on other individual characteristics and conditions. In particular:

$$A_i = x_{0i}\alpha_0 + \varepsilon_{0i} \tag{1}$$

where x_{0i} is a vector of variables, including past consumption behaviour, that are hypothesized to affect the expected fixed costs of quitting, α_0 is the corresponding vector of parameters, and ε_{0i} is a random disturbance.

The expected benefits of quitting represent the utility derived by giving up smoking and reflect the health, financial and social benefits of habit cessation. Formally, we assume that the expected benefits of quitting are given by:

$$B_i = B(y_i) = B(g^{-1}(y^*_{2i})) \tag{2}$$

where y_i is the observed level of consumption and y_{2i}^* is a latent variable that characterizes the individual's demand for cigarettes, such that $y_{2i}^* = g(y_i)$ (where $g(\cdot)$ is an increasing transformation). It seems to be reasonable to assume that the benefits of quitting depend on how much the individual would have smoked otherwise and that they increase, at a decreasing rate, as the desired and observed level of cigarette consumption rises ($B'(\cdot) > 0$, $B''(\cdot) < 0$).

The decision to attempt to stop smoking will depend on the expected net benefit of quitting ($B - A$), that is on whether the benefits exceed the fixed costs of nicotine dependence (Jones, 1994). The condition for an individual to continue smoking and not attempting to quit can be then written as:

$$y_{1i}^* = A_i - B_i = 0 \tag{3}$$

Specifying individual's demand for cigarettes as:

$$y_{2i}^* = x_{2i}\alpha_2 + \varepsilon_{2i} \tag{4}$$

where x_2 is a vector of variables determining the demand for cigarettes, which may include the addiction effects of tolerance and reinforcement, α_2 is the corresponding vector of parameters, and ε_{2i} is a random disturbance. Assuming for tractability that $B(\cdot)$ is the identity, we obtain:

$$y_{1i}^* = x_{0i}\alpha_0 + \varepsilon_{0i} - x_{2i}\alpha_2 - \varepsilon_{2i} \tag{5}$$

or, equivalently:

$$y_{1i}^* = x_{1i}\alpha_1 + \varepsilon_{1i} \tag{6}$$

where x_1 is the union of x_0 and x_2, α_1 is the corresponding parameter vector, where for variables that appear in both x_0 and x_2 we have $\alpha_1 = \alpha_0 - \alpha_2$ and $\varepsilon_1 = \varepsilon_0 - \varepsilon_2$.

Equation (6) gives the first hurdle that an individual must overcome to be observed with a positive level of consumption. In line with a double-hurdle specification, we also allow for the latent variable y_{2i}^* to generate zero observations. Then, the observed level of cigarette consumption is equal to:

$$y_i = \begin{cases} g^{-1}(y_{2i}^*) & \text{if } y_{1i}^* > 0 \text{ and } g^{-1}(y_{2i}^*) = g^{-1}(x_{2i}\alpha_2 + \varepsilon_{2i}) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

It should be noticed that for variables that have no (or small) influence on the fixed costs of quitting ($\alpha_0 \cong 0$) we would expect equal and opposite effects on the two decisions ($\alpha_1 \cong -\alpha_2$) (Yen and Jones, 1996).

Given the specification in equations (6) and (7) and assuming that ε_{0i} and ε_{2i} have zero mean, constant variances ($\sigma^2_{\varepsilon 0}$ and $\sigma^2_{\varepsilon 2}$), and covariance $\sigma_{\varepsilon 02}$, the correlation coefficient between ε_{1i} and ε_{2i} is:

$$\rho_{\varepsilon 12} = \frac{\sigma_{\varepsilon 02} - \sigma^2_{\varepsilon 2}}{\sigma_{\varepsilon 2}\sqrt{\sigma^2_{\varepsilon 0} + \sigma^2_{\varepsilon 2} - 2\sigma_{\varepsilon 02}}} \tag{8}$$

which depends on the degree of correlation between the unobservables of the fixed costs of quitting (A) and the desired level of consumption (y^*_{2i}) and suggests that it would not be appropriate to *a priori* assume independence between the error terms of the equations in the double-hurdle model.

2.2 Social interactions: within-household peer effects

The importance of social interactions on smoking has been recently emphasized by a growing body of both economic and public health literatures (Ary and Biglan, 1988; Krauth, 2005). Social interactions are widely regarded as important determinants of many behavioural and economic outcomes, based on the idea that the utility that an individual receives from doing a certain activity depends on the actions of the other individuals in the person's reference group (Becker, 1996). The net benefit of consuming a given good will then increase with the consumption of the same good of other individuals. Therefore, the point at issue is to verify whether the average behaviour in a group affects the behaviour of the individuals in that particular group (Manski, 1993, 1995; Becker, 1996; Brock and Durlauf, 2001).

Several studies have shown that social interaction matters in the decision to quit smoking (Jones, 1994; Yen and Jones, 1996). Hence, we extend the model outlined in the previous Section 2.1, by including a measure of social interaction in both smoking participation and cigarette demand equations. In particular, we focus on the within-household interactions only, analyzing the effects exerted on individual's attitude toward tobacco consumption by the smoking behaviour of the other components of the family (Auld, 2005).

The model of quitting and smoking with household-peer effects is similar to standard models with social interaction proposed in the literature (Brock and Durlauf, 2001; Powell *et al.*, 2005; Krauth, 2005). Each individual is a member of a peer group (family) and is assumed to be influenced symmetrically by each other member and there are no cross-group influences. Individuals choose whether to smoke, and the level of cigarette consumption, ($y_i > 0$) or not ($y_i = 0$): the relative utility from each choice is a function of the individual and group's observed characteristics, the average choice in the reference group, and a random utility term. Indexing groups by g (with n_g denoting the dimension of the group) and individuals by i, individual's utility function $u_{i,g}(y_i)$ satisfies the following condition:

$$u_{i,g}(y_i) - u_{i,g}(0) = F\left[x_{1i,g}, z_{1g}, \overline{s}_{i,g}, \varepsilon_{1i,g}; x_{2i,g}, z_{2g}, \overline{c}_{i,g}, \varepsilon_{2i,g}\right] \tag{9}$$

where $x_{1i,g}$ and $x_{2i,g}$ are individual-level exogenous variables and z_{1g} and z_{2g} are group-level variables, affecting participation and consumption respectively, $\overline{s}_{i,g}$ is the fraction of the other smokers within the group:

$$\overline{s}_{i,g} = \frac{1}{n_g - 1} \sum_{j \neq i} s_{j,g}\ , \tag{10}$$

$\overline{c}_{i,g}$ is the number of cigarette consumed by the other group members:

$$\overline{c}_{i,g} = \sum_{j \neq i} c_{j,g} \tag{11}$$

and $\varepsilon_{1i,g}$ and $\varepsilon_{2i,g}$ are unobserved individual-level terms

The model thus incorporate what Manski (1993) defines "endogenous peer effects" ($\overline{s}_{i,g}$ and $\overline{c}_{i,g}$) into a double-hurdle limited dependent variable model of smoking. It is worth noting that, following the approach commonly adopted in the empirical literature on peer effects (Sacerdote, 2001; Gaviria and Raphael, 2001; Krauth, 2005), the model does not include the "contextual effects", i.e. the effects on the net utility benefit from smoking exerted by the characteristics (as opposed to smoking behaviour) of individual's group, given the difficulties of simultaneously estimate "endogenous peer effects" and "contextual effects" (Manski, 1993).

3 Econometric specification

The empirical counterpart to the model outlined in the previous section is the double-hurdle model proposed by Jones (1989). The main feature of the double-hurdle approach is that participation and consumption decisions are assumed to stem from two separate individual choices and the determinants of the two decisions are allowed to differ. In order to specify the model, we assume the error terms to be distributed as a bivariate normal and we normalize the distribution so that $\sigma^2_{\varepsilon 1} = 1$. Formally, the two equations of the model, corresponding to equation (4) and (6) can be written as[1]:

$$y^*_{1i} = x'_{1i}\beta_1 + u_{1i} \quad (12)$$

$$y^*_{2i} = x'_{2i}\beta_2 + u_{2i} \quad (13)$$

where:

$$(u, v) \sim BVN(0, \Sigma) \quad \text{with} \quad \Sigma = \begin{bmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma^2 \end{bmatrix}$$

Following Yen and Jones (1996), for the increasing transformation $g(\cdot)$ in equation (2) we use a Box-Cox transformation, which introduces more flexibility and allows to recover the normality assumption when dealing with high skewed data. The Box-Cox transformation of the dependent variable gives:

$$y^*_{2i} = \begin{cases} (y_i^\lambda - 1)/\lambda & \text{if} \quad \lambda > 0 \\ \log(y_i) & \text{if} \quad \lambda = 0 \end{cases} \quad (14)$$

where λ is an unknown parameter such that $0 \le \lambda \le 1$. Given this transformation, the censoring mechanism implies the following relationship between the dependent variable (y_i) and the latent variables (y^*_{2i} and y^*_{1i}):

$$y^*_{2i} = \begin{cases} (y_i^\lambda - 1)/\lambda & \text{if} \quad y^*_{2i} > -1/\lambda \ \text{and} \ y^*_{1i} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

where y^*_{1i} and y^*_{2i} are defined as in (12) and (13).

This specification relaxes the normality assumption on the conditional distribution of y_i and still allows stochastic dependence between the error terms. The likelihood functions for the dependent Box-Cox double-hurdle model can be written as:

---

[1] The parameter vectors β_1 and β_2 are introduced to reflect the normalization of the variance.

$$L_1 = \prod_0 \left[1 - \Phi\left(x'_{1i}\beta_1, \frac{x'_{2i}\beta_2 + 1/\lambda}{\sigma}, \rho \right) \right]$$

$$\times \prod_+ \left[\Phi\left(\frac{x'_{1i}\beta_1 + (\rho/\sigma)\left[(y_i^\lambda - 1)/\lambda - x'_{2i}\beta_2 \right]}{\sqrt{1-\rho^2}} \right) y_i^{(\lambda-1)} \frac{1}{\sigma} \phi\left(\frac{(y_i^\lambda - 1)/\lambda - x'_{2i}\beta_2}{\sigma} \right) \right] \quad (16)$$

where Φ denotes the standard normal CDF (univariate or multivariate), ϕ is the univariate standard normal PDF, $\rho = \sigma_{12}/\sigma$ and 0 and + denote zero and positive consumption, respectively.

The relevance of model (16) rests on the fact that it is a general model that nests other double-hurdle specifications and also encompasses a wide range of limited dependent variable models (Yen, 1993; Jones and Yen, 2000). In particular, by placing the appropriate restrictions on λ and ρ, three main nested specifications of interest can be obtained.

1) $\rho = 0$ (i.e. $\sigma_{12} = 0$), which gives the Box-Cox double-hurdle model with independent error terms (Jones and Yen, 2000):

$$L_2 = \prod_0 \left[1 - \Phi(x'_{1i}\beta_1)\Phi\left(\frac{x'_{2i}\beta_2 + 1/\lambda)}{\sigma} \right) \right] \prod_+ \left[\Phi(x'_{1i}\beta_1) y_i^{(\lambda-1)} \frac{1}{\sigma} \phi\left(\frac{(y_i^\lambda - 1)/\lambda - x'_{2i}\beta_2}{\sigma} \right) \right] \quad (17)$$

2) $\lambda = 1$, which gives the standard double-hurdle model with dependent errors (Jones, 1989, 1992):

$$L_3 = \prod_0 [1 - \Phi(x'_{1i}\beta_1, x'_{2i}\beta_2, \rho)] \prod_+ \left[\Phi\left(\left[z'_i\alpha + \frac{\rho}{\sigma}(y_i - x'_i\beta) \right] \Big/ \sqrt{1-\rho^2} \right) \frac{1}{\sigma} \phi\left((y_i - x'_i\beta)/\sigma \right) \right] \quad (18)$$

that can be further restricted by assuming independence of the errors ($\rho = 0$) to obtain the independent double-hurdle model, which has been widely used in empirical analysis of tobacco consumption (Blaylock and Blisard, 1992; Garcia and Labeaga, 1996). Moreover, the independent double-hurdle model nests the standard Tobit model (Tobin, 1958), allowing to test for the relevance of a bivariate generalization (Garcia and Labeaga, 1996).

3) $\lambda = 0$, which gives the generalized Tobit model (Heckman, 1979) with $\log(y_i)$ as the dependent variable of the second hurdle:

$$L_4 = \prod_0 \left[1 - \Phi\left(x'_{1i}\beta_1\right)\right]$$
$$\times \prod_+ \left[\Phi\left(\frac{x'_{1i}\beta_1 + (\rho/\sigma)\left[\log(y_i) - x'_{2i}\beta_2\right]}{\sqrt{1-\rho^2}} \right) y_i^{\lambda} \frac{1}{\sigma} \phi\left(\frac{\log(y_i) - x'_{2i}\beta_2}{\sigma} \right) \right] \qquad (19)$$

By imposing $\rho = 0$, model (19) reduces to the "Two-part" Model (Blaylock and Blisard, 1992), that corresponds to a probit equation for participation and an OLS of $\log(y_i)$ on x_{2i} for the positive observations.

As highlighted in Section 2.2, controls for social interactions are included in both the equations of the double-hurdle model. In this analysis, we define two within-household peer smoking measures. The first peer measure is constructed as the prevalence of smokers in the individual's household not including the given individual in calculation (*peer participation rate*). In line with previous works on social interaction (Powell *et al.*, 2005; Krauth, 2005), we assume that individual's decision to continue smoking (i.e. the first hurdle of the model) is influenced by the average prevalence of smoking among all other respondents in his(/her) family. The second peer measure considered is the average number of cigarettes consumed by the other smokers of the household (*peer consumption*). This variable is intended to capture social aspects of cigarette consumption and to measure the effects of intra-household consumption externalities on individual smoking behaviour.

The main issue related to the inclusion of peer-group effects in the double-hurdle specification is the potential endogeneity of the variables measuring social interactions (Auld, 2005). However, it is not feasible to expand the model to allow them to be simultaneously determined, owing to both data limitations (further instruments would be required) and computational feasibility. Since either conditioning on or excluding endogenous covariates may be problematic, in order to check the robustness of the results to the inclusion of these potentially endogenous regressors the model is estimated including and not including the within-household peer smoking measures. Moreover, as highlighted by Manski (1993) and Krauth (2005), the failure to account for the simultaneity of choice among peers leads to overestimate the actual strength of peer influence. For this reason, the dimension of the estimated measures of social interaction should be cautiously evaluated.

4 Data and empirical results

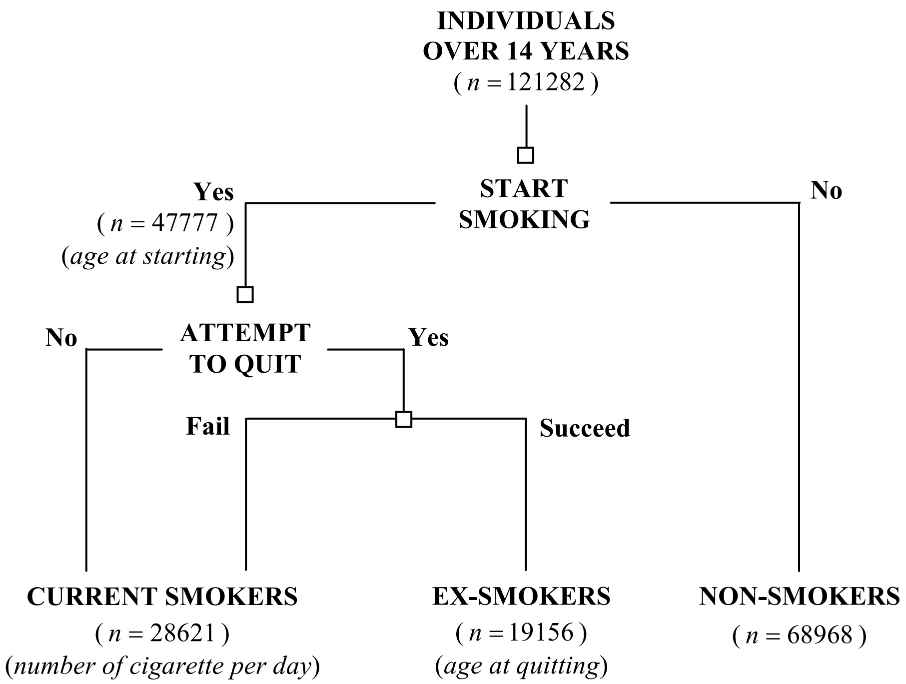
4.1 The ISTAT survey on "Health status and use of health services"

The dataset used in this study is the multiscope survey on "Health status and use of health services" (HSHS, from now onwards), conducted by the Italian National Statistical Institute (ISTAT) between September 1999 and June 2000. The HSHS survey is a representative cross-section sample of the Italian population and provides detailed information on demographic and social characteristics, health conditions, and utilization of health institutions of 140011 individuals living in private households.

For the aims of the present study, the relevance of this survey lies in the detailed section devoted to the analysis of current and past smoking habits of individuals aged 14 and over. In particular, we focus on the determinants of individual smoking behaviour, in terms of quitting and cigarette consumption decisions. Moreover, since the model presented in Section 2 highlights the importance of addiction in quitting decision, we need data on past consumption behaviour and on the history of the smoking habit. In this respect, the HSHS survey is the most complete source of data on cigarette consumption available in Italy. Contrary to the ISTAT "Italian Household Budget Survey", in fact, it provides information on individuals rather than on households, allowing a thorough analysis of "real" socio-demographic and gender effects, without approximating them with the characteristics of the household's head. Moreover, individual data enable to evaluate within-household social interactions and consumption externalities. The HSHS survey, just like the other ISTAT multiscope survey on "Everyday Life Aspects", also provides information that not only distinguishes between "smokers" and "non-smokers", but separates non-smokers into those who have never smoked and those who class themselves as ex-smokers. This enables to extend the analysis to distinguish between starting and quitting decisions. Information on cigarette consumption behaviour is completed by the age at which the individual has started and/or quitted smoking, that allows to obtain a measure of the duration of the smoking habit[2].

[2] This information is not available in the ISTAT "Everyday Life Aspects" survey.

Figure 1 – Smoking behaviour of the individuals in the sample



***Notes**:* occasional (current and ex) smokers are excluded from the sample.

In order to estimate the double-hurdle models in Section 3, the sample has to be restricted to current and ex-smokers, so that participation corresponds to the decision to whether or not quit smoking (Jones, 1989; Yen and Jones, 1996). The final sample consists of 47777 individuals, 60 per cent of whom class themselves as current smokers[3]. Figure 1 summarizes all the aspects of the attitude towards smoking based on the information derived from the HSHS survey and shows how observed individuals can be divided into current smokers, ex-smokers and those who have never smoked.

4.2 Variables and model specification

In this study, the number of cigarettes smoked per day (*N_Cig*) is used as the dependent variable of the consumption equation. This is a volume based variable typical of health surveys that, contrary to an expenditure-based measure, does not control for differences

[3] We exclude from the sample occasional (current and ex) smokers.

in prices and quality of cigarettes consumed. However, being a measure of "typical consumption", it is not influenced by undetected infrequency of purchase problems that often arise in expenditure surveys (Blundell and Meghir, 1987; Yen, 2005a), enabling us to analyze the decision to quit smoking and not participating to cigarette consumption as stemming from conscious abstentions or corner solutions only. It should be also noted that the participation variable (*D*) measures the prevalence of smoking and so it reflects the number of individuals who have quit up to the time of the survey. As in Jones (1994) and Yen and Jones (1996), the results should therefore be interpreted in terms of the stock of individuals who have quit rather than the flow of new quits over a specific period.

Individual cigarette consumption is assumed to be expressible as a linear combination of explanatory variables that are hypothesized to encompass the determinants of the quitting decision and the level of cigarette consumption. Together with measures of addiction and social interactions, they include information on health conditions and demographic and socio-economic status, which are hypothesized to separately affect quitting and smoking decisions. Given that the HSHS survey is a single cross-section it is not possible to identify the impact of prices on cigarette consumption.

One estimation issue in double-hurdle models concerns the choice of the regressors for participation and consumption equations. In particular, since including the same set of regressors in each hurdle makes the parameters identification difficult, exclusion restrictions must be imposed[4]. For this reason, before presenting estimation results, a discussion of the explanatory variables included in the model is necessary.

In this study, both the probability of quitting and consumption level are assumed to be influenced by the duration of the smoking habit (*Smoking_Habit*), demographic characteristics (*Age* and *Male*), zone of residence (*North* and *South*), marital status (*Single*), health and physical characterisitics (*Chronic_Illness*, *Physical_Act*, *BMI*), education level (*Education*) and social and occupational status (*Professional*, *Whitecollar, Unemployed*, *Student*, *Retired*) of the individual.

Given the aims of the study, in order to account for individual's past smoking behaviour and to control for the fixed costs associated with addiction, a continuous variable indicating the duration (measured in years) of the smoking habit (*Smoking_Habit*) has been included in the model. Even if the cross-sectional nature of the

[4] In a pre-estimation phase, we started with a specification that included all explanatory variables in both hurdles and then we gradually excluded insignificant variables, giving identification higher reliability.

HSHS survey prevents the analysis of and consumption dynamics, which would require the use of (genuine or pseudo) panel data as in Jimenez-Martin *et al*. (1998), Labeaga (1999) and Jones and Labeaga (2003), this variable represents a measure of the addictive stock of past cigarette consumption[5], that may rise the cost of quitting (Suranovic *et al.*, 1999; Jones, 1999), and allows to analyze the effect of past smoking behaviour.

Concerning demographic factors, age is considered to assess the existence of a lifecycle patterns in cigarette consumption and to verify whether age-related health problems affect smoking behaviours, while the gender dummy variable *Male* (equal to 1 if the individual is a man) is included to account for gender-differentiated smoking habits (Yen, 2005a).

In previous empirical studies (Blaylock and Blisard, 1992; Yen, 2005b; Garcia and Labeaga, 1996; Yen and Jones, 1996), education level has been found to significantly affect smoking behaviour, with better educated individuals moderating cigarette consumption as they are more aware of the health risks connected with consumption of unhealthy.

Social and working status variables status (*Professional*, *Whitecollar, Unemployed*, *Student*, *Retired*) may help in explaining how the prevalence of smoking and levels of cigarette consumption vary among different social groups, reflecting differences in attitudes towards health and time preferences.

Three health and physical status variables are assumed to affect both quitting probability and consumption. Specifically, we use a binary variable for whether the individual suffers, or has suffered in the past, from a or long-standing illness or disability (*Chronic_Illness*). Finally, we include a clinical measure of the general physique by the Quetelet's body mass index (*BMI*) and we assume that physical activity (*Physical_Act*) may reflect health awareness attitudes developed over the lifetime (Blaylock and Blisard, 1992, 1993) and be an important factor in moderating consumption.

A self-evaluated measure of current health status (*Health*) has been included in the consumption equation only. This dummy variable (which equals 1 if self-rated health status is fair/poor) is constructed on the basis of the standard (excellent/good/fair/poor) scale for self-assessed health reported in the HSHS. As pointed out by Blaylock and Blisard (1992) and Shmueli (1996), current health status is not a good predictor of smoking participation, but may be associates with the decision concerning the level of cigarette consumption. In particular, in our study the health status at the time of

[5] This is a plausible hypothesis given that only habitual smokers, who currently smoke or were used to smoke more than one cigarette per day, are considered.

cessation is unobserved and relating the “stock” of quitters to their present health is misleading. The interval between the time at which the individual is observed and the time of quitting, in fact, causes an unobserved heterogeneity that masks the true relationship between health and smoking cessation (Shmueli, 1996; Jones, 1996).

Specific variables accounting for economic conditions have been introduced in the consumption equation. The HSHS survey does not provide any information on individual income and information on economic conditions are recovered by means of different explicative factors. In particular, we consider a dummy variable (*Ec_Resources*) that account for individual’s self-rating of the economic resources of his(/her) household (which is equal to one is economic resources are considered insufficient) and the number of earners within the household (*N_Earners*). Following the suggestions of Atkinson *et al.* (1984) and Jones (1989), a variable indicating whether the household lives in a home that is owned or being bought (*OwnerOcc*) is included as a proxy for wealth and economic stability. However, the use of proxies for income leads to measurement errors and so results for the socio-economic variables should be viewed with caution.

Finally, as previously discussed in Sections 2.2 and 3, two different measures of social interactions in the two hurdles. In line with binary endogenous social interaction models (Brock and Durlauf, 2001; Krauth, 2005), a variable indicating the smoking participation rate of the other members of the household (*Peer_Part*) is included in the participation equation. On the other hand, the consumption equation is extended to account for intra-household consumption externalities on individual smoking behaviour, by means of the average number of cigarette smoked by the other members of the household (*Peer_Cons*).

Detailed description and sample statistics for all the variables are presented in Table 1. The sample ($n = 47777$) is stratified between men ($n = 31912$) and women ($n = 15865$). Of these, 17947 men (56% of the total) and 10674 women (67%) are current smokers, with an average consumption of cigarette smoked per day of 9.20 and 7.99, respectively.

Table 1 – Variable definitions and descriptive statistics

Variable	Definition	Mean[a]		
		Entire Sample (n=47777)	*Men (n=31912)*	*Women (n=15865)*
DEPENDENT VARIABLES				
D	Equals 1 if a smoker	0.60	0.56	0.67
N_CIG	Number of cigarettes per day	8.80 (10.12)	9.20 (10.82)	7.99 (8.47)
EXPLANATORY VARIABLES (CONTINUOUS)				
SMOKING_HABIT	Smoking habit duration in years	23.47 (14.31)	24.92 (14.84)	20.56 (12.68)
PEER_CONS	Number of cigarette consumed per day by the other members of the household	12.06 (15.68)	9.79 (14.46)	16.62 (16.98)
PEER_PART	Smoking participation rate of the other members of the household	0.17 (0.22)	0.15 (0.21)	0.21 (0.24)
BMI	Quetelet's Body Mass Index[b]	24.89 (3.81)	25.63 (3.50)	23.40 (3.98)
AGE	Age in years	47.10 (16.94)	48.66 (17.46)	43.95 (15.39)
EDUCATION	Years of formal education	8.90 (4.14)	8.56 (4.13)	9.59 (4.07)
N_EARNERS	Number of earners within the household	0.93 (0.81)	0.89 (0.79)	1.01 (0.84)
EXPLANATORY VARIABLES (BINARY)				
HEALTH	Equals 1 if self-rated health status is fair/poor	0.46	0.45	0.46
CHRONIC_ILLNESS	Equals 1 if individual is affected by cardiac and respiratory chronic illness	0.24	0.27	0.20
PHYSICAL_ACT	Equals 1 if individual does physical activity (at least once a week)	0.71	0.71	0.70
MALE	Equals 1 if male	0.67	–	–
SINGLE	Equals 1 if single adult household (with or without young children)	0.15	0.13	0.20
CHILDREN013	Equals 1 if any child aged 0-13 is present within the household	0.27	0.25	0.31
PROFESSIONAL	Equals 1 if professional	0.03	0.04	0.02
WHITECOLLAR	Equals 1 if white-collar	0.23	0.20	0.28
UNEMPLOYED	Equals 1 if unemployed	0.08	0.07	0.09
STUDENT	Equals 1 if student	0.04	0.03	0.05
RETIRED	Equals 1 if retired	0.25	0.31	0.14
NORTH	Equals 1 if individual resides in the North	0.42	0.40	0.46
SOUTH	Equals 1 if individual resides in the South	0.39	0.42	0.33
OWNER_OCC	Equals 1 if household own its home	0.74	0.75	0.71
EC_RESOURCES	Equals 1 if self-reported household economic resources are considered insufficient	0.28	0.29	0.28
QUARTER1	Equals 1 if observed in September 1999	0.26	0.26	0.26
QUARTER2	Equals 1 if observed in December 1999	0.25	0.25	0.25
QUARTER3	Equals 1 if observed in March 2000	0.25	0.25	0.24

Notes: [a] Standard deviations in parentheses
[b] Quetelet's Body Mass Index = (weight in kg)/(height in metres)^2

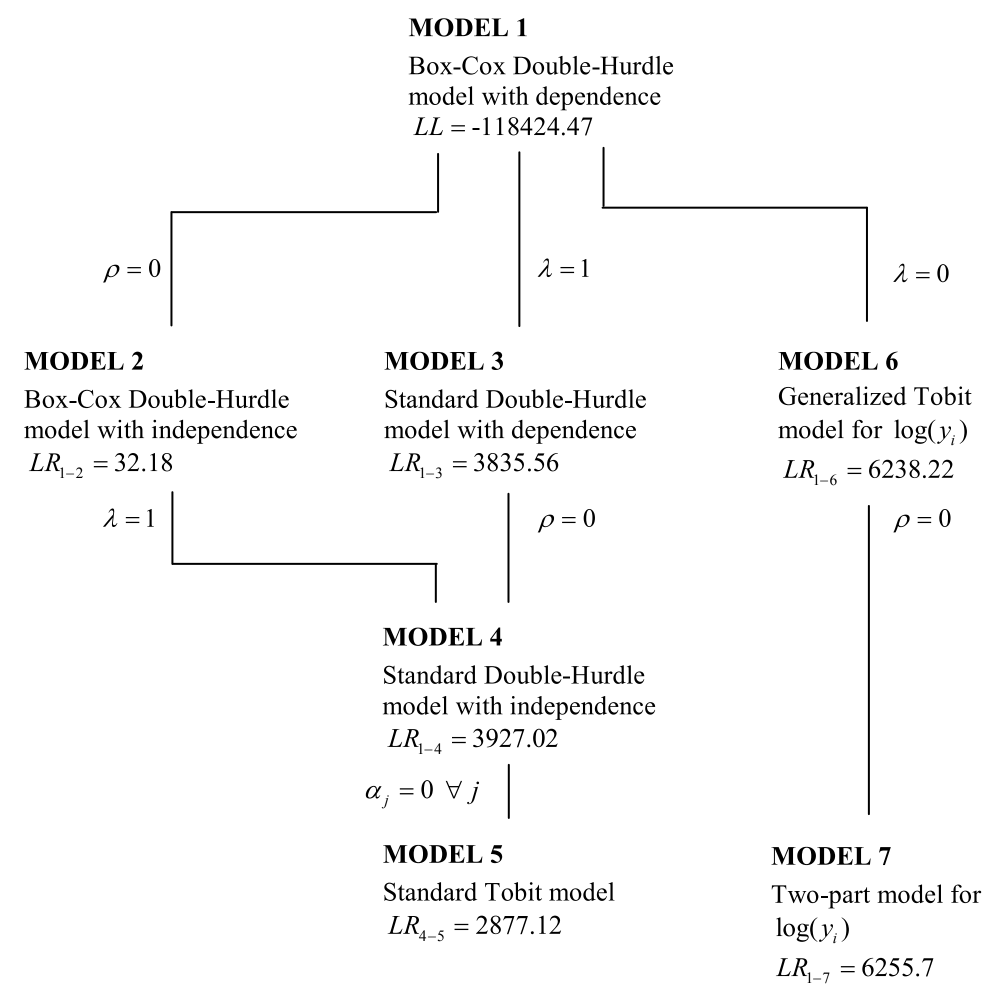
All the models discussed in Section 3 have been estimated by maximizing the logarithm of the likelihood functions (16) and then placing the appropriate restrictions to obtain the relevant nested specification (17), (18) and (19). The likelihood ratio tests, together with the nesting structure that links all the models, are presented in Figure 2. It is worth noticing that the validity of the LR tests for nested specifications rests on the assumption that the general model (16) is not misspecified. For these reason, a preliminary LR test for homoscedasticity have been performed on the generalized double-hurdle specification; the results for the first and second hurdle (LR statistics equal to 1.143 ($p-value=0.3350$) and 1.657 ($p-value=0.1271$), respectively) indicates that the chosen specification is homoscedastic. Moreover, we carried out a RESET-type misspecification test using the second, third and fourth powers of the fitted values for the consumption equation as extra regressors. The corresponding LR statistic is equal to 0.34 ($p-value=0.7964$), providing no evidence of misspecification.

All the restricted specification are rejected, each with *p-value* lower than 0.001. From these results emerge a clear indication of the opportunity of a flexible specification that allows to separately analyze the determinants of participation and consumption decisions, relaxes the hypothesis of normality and accounts for dependence between the error terms of the two hurdles.

4.3 Parameter estimates and elasticities

Maximum-likelihood estimates of the Box-Cox double-hurdle model (16), with and without social interactions, are presented in Table 2. The variables measuring the smoking behaviour of the respondents' family, as previously discussed in Section 3, are potentially endogenous and a check on the robustness of the results to the inclusion of within-household interactions is therefore necessary (Auld, 2005). As it can be noted from the inspection of the Table, when the household-peer effects are not held constant, all the estimated parameters (including the Box-Cox parameter λ) remains stable. Results are very similar with and without conditioning on the behaviour of the other components of the respondent's household, with little variation in the estimates and no sign changes. Moreover, both the measures of social interaction considered are found to

Figure 2 – Model specification: likelihood-ratio tests for nested specifications



Adapted from Yen and Jones (1996)
***Notes**: LL = log-likelihood, LR = LR test statistic*

significantly affect individual smoking behaviour[6]. Household smoking participation rate significantly increases the probability of remaining a smoker, while the number of cigarettes consumed by the other smokers within the household increases the level of individual consumption, revealing the existence of significant consumption externalities (Maurer and Meier, 2005).

[6] Given the potential endogeneity of the peer participation variable, as pointed out by Krauth (2005), the corresponding parameter in the participation equation may provide a measure that overstates the strength of the real peer influence on smoking participation.

Table 2 – Box-Cox double-hurdle estimates with and without social interactions

Variable	With social interactions		Without social interactions	
	Participation	*Consumption*	*Participation*	*Consumption*
SMOKING_HABIT	0.1110*** (0.002)	0.1035*** (0.005)	0.1122*** (0.002)	0.1122*** (0.005)
PEER_CONS	–	0.0217*** (0.001)	–	–
PEER_PART	0.9059*** (0.036)	–	–	–
HEALTH	–	0.0732*** (0.028)	–	0.0877*** (0.029)
CHRONIC_ILLNESS	-0.2252*** (0.019)	-0.0908*** (0.039)	-0.2172*** (0.019)	-0.0795** (0.040)
PHYSICAL_ACT	-0.0748*** (0.016)	-0.2708*** (0.028)	-0.0794*** (0.016)	-0.2861*** (0.029)
BMI	-0.0250*** (0.002)	0.0320*** (0.004)	-0.0249*** (0.002)	0.0301*** (0.004)
MALE	-0.3036*** (0.017)	0.9483*** (0.036)	-0.3517*** (0.017)	0.8140*** (0.036)
AGE	-0.1102*** (0.002)	-0.0804*** (0.005)	-0.1125*** (0.002)	-0.0904*** (0.005)
SINGLE	0.2740*** (0.023)	0.4442*** (0.037)	0.1787*** (0.022)	0.3030*** (0.037)
CHILDREN013	-0.2135*** (0.017)	–	-0.2187*** (0.017)	–
EDUCATION	-0.0077*** (0.002)	-0.0129*** (0.004)	-0.0086*** (0.002)	-0.0167*** (0.004)
PROFESSIONAL	0.0138 (0.040)	0.1334* (0.080)	0.0087 (0.040)	0.1143 (0.082)
WHITECOLLAR	-0.0535*** (0.019)	0.0375 (0.034)	-0.0589*** (0.019)	-0.0101 (0.035)
UNEMPLOYED	0.0687*** (0.029)	-0.2854*** (0.046)	0.0847*** (0.029)	-0.0320 (0.045)
STUDENT	0.0378 (0.045)	-1.1951*** (0.064)	0.0531 (0.045)	-0.9475*** (0.063)
RETIRED	-0.2142*** (0.025)	-0.7458*** (0.055)	-0.2299*** (0.025)	-0.5890*** (0.055)
NORTH	-0.0329* (0.020)	-0.0696** (0.034)	-0.0417*** (0.020)	-0.1069*** (0.035)
SOUTH	0.0951*** (0.020)	0.1071*** (0.034)	0.0800*** (0.020)	0.1143*** (0.035)
OWNER_OCC	–	-0.2455*** (0.028)	–	-0.2420*** (0.029)
N_EARNERS	–	-0.0997*** (0.021)	–	-0.1178*** (0.019)
EC_RESOURCES	–	0.0782*** (0.028)	–	0.1261*** (0.029)
QUARTER1	–	-0.0364 (0.035)	–	-0.0324 (0.035)
QUARTER2	–	-0.0809*** (0.035)	–	-0.0869*** (0.035)
QUARTER3	–	-0.0946*** (0.035)	–	-0.0986*** (0.036)
CONSTANT	3.6974*** (0.073)	4.7844*** (0.157)	3.9874*** (0.071)	5.1868*** (0.170)
σ	2.0564*** (0.050)		2.1130*** (0.057)	
ρ	-0.1931*** (0.023)		-0.2518*** (0.022)	
λ	0.4644*** (0.010)		0.4691*** (0.011)	
Log Likelihood	-118424.47		-119058	

Notes: robust standard errors are reported in parentheses.
***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Given these results, in the rest of the discussion we focus the attention on the results of the Box-Cox model with social interactions. The Box-Cox parameter λ is equal to 0.4644 and is significantly different from both zero and one, leading to reject both the standard double-hurdle model and the generalized Tobit with $\log(y_i)$ as dependent variable[7]. Moreover, the parameter ρ is significantly different from zero at the 1% level, supporting the hypothesis, suggested by the fixed costs model, of strong correlation between the error terms of the two hurdles.

In the full Box-Cox double-hurdle model the interpretation of the estimated effects of the explanatory variables on the probability of quitting and on the observed level of cigarette consumption is complicated by the dependence between the error terms of the two hurdles and by the nonlinear transformation of the dependent variable. Thus, in order to obtain a more intuitive interpretation of the effects of explanatory variables the elasticities of participation, conditional and unconditional mean are calculated[8], following the decomposition proposed by Yen and Jones (2000)[9], and presented in Table 3.

Analysing the different effects of the continuous variables for consumption decision, we find that *N_Earners* has a negative impact on cigarette consumption. This variable, proxing the incremental effect of household's income on the level of consumption, is in line with the findings of Garcia and Labeaga (1996) and Aristei and Pieroni (2007) and imply that, as for most goods in consumer choice models, cigarette consumption rises as household income increases, but at a decreasing rate.

The estimated elasticities with respect to age suggest that older smokers are less likely to smoke and, conditional on smoking, consume less tobacco than younger. Moreover, educated smokers have a lower level of cigarette consumption supporting the hypothesis that educated individuals are more aware of the health risks associated with smoking. These results are fully consistent with previous studies on US (Yen, 1999), UK (Jones, 1989; Yen and Jones, 1996) and Spanish (Garcia and Labeaga, 1996) data.

[7] The estimated parameter is close to the square root transformation (even if it is significantly different from 0.5). As pointed out by Yen and Jones (1996) and Yen (1999), this may suggest the appropriateness of using the Negative Binomial Hurdle Model (Pohlmeier and Ulrich, 1995), interpreting the number of cigarettes smoked as a count variable. However, the Negbin Hurdle Model does not allow for dependence between participation and consumption equations, which is implied by the fixed cost model and is found to be significant in the estimation on model (16).

[8] For dummy explanatory variables, the effects on probability, conditional level and unconditional level are computed as the difference in these components when the value of the dummy shifts from zero to one, holding all the other regressors constant.

[9] Details on the derivation of the elasticities for the Box-Cox Double Hurdle model with dependent errors can be found in Jones and Yen (2000).

Table 3 – Elasticities with respect to continuous variables and effects of binary variables

Variables	*Probability*	*Conditional level*	*Unconditional level*
Continuous variables			
SMOKING_HABIT	1.5301*** (0.0709)	0.3193*** (0.0153)	1.8494*** (0.0818)
PEER_CONS	–	0.0344*** (0.0017)	0.0344*** (0.0017)
PEER_PART	0.1277*** (0.0103)	–	0.1277*** (0.0103)
BMI	-0.3916*** (0.0902)	0.1048*** (0.0127)	-0.2864*** (0.1207)
AGE	-3.2481*** (0.1391)	-0.4979*** (0.0285)	-3.7460*** (0.1583)
EDUCATION	-0.0493 (0.0138)	-0.0152*** (0.0047)	-0.0645*** (0.0231)
N_EARNERS	–	-0.0122*** (0.0026)	-0.0122*** (0.0026)
Discrete variables			
HEALTH	0.0000	0.2923***	0.1706*
CHRONIC_ILLNESS	-0.0886***	-0.1491**	-1.2377***
PHYSICAL_ACT	-0.0292**	-1.0177***	-0.9861***
MALE	-0.1159***	4.0000***	0.9614***
SINGLE	0.1042***	1.5499***	2.3754***
CHILDREN013	-0.0839***	0.2018*	-0.9857***
PROFESSIONAL	0.0054	0.5231*	0.3780*
WHITECOLLAR	-0.0209***	0.1999	-0.1588*
UNEMPLOYED	0.0264**	-1.1885***	-0.3733**
STUDENT	0.0132	-4.5608***	-2.5380***
RETIRED	-0.0846***	-2.7330***	-2.5830***
NORTH	-0.0129*	-0.2475**	-0.3119**
SOUTH	0.0371**	0.3401**	0.6849***
OWNER_OCC	-0.0001	-0.9854***	-0.5752***
EC_RESOURCES	0.0000	0.3127***	0.1826*

Notes: Asymptotic standard errors of estimated elasticities are reported in round brackets.
Asterisks indicate levels of significance: *** = 0.01, ** = 0.05 and * = 0.10.

The level of cigarettes consumption of the other smokers within the household (*Peer_Cons*) positively affects individual consumption, revealing the existence of significant externalities on smoking behaviour. This finding is confirmed in the probability estimation by the significant influence of household's participation rate (*Peer_part*) on the probability of quitting. Individuals with smokers within their households are less likely to have quit, confirming the results of Jones (1994).

As expected by the findings in previous works (Jones, 1999; Jones and Labeaga, 2003), also for Italian smokers habit significantly affects the consumption level and quitting decision. In particular, the estimated parameter for the probability to quit is very large supporting the hypothesis of high fixed costs associated with addictive behaviour.

The effect of physical status, proxied by the body mass index, shows that individuals with a higher *BMI* are less likely to continue smoking, but tend to smoke more cigarettes. The analysis of the effects of binary variables for health status gives further emphasis to the evidence highlighted in the discussion of estimated parameters of the *BMI*. It is worth noticing that *Chronic_Illness* and *Physical_Act* simultaneously reduce the probability to continue smoking and the level of consumption showing, on the one hand, that the status of current or past long standing illness requires to cut smoking and, on the other hand, that individuals more aware to health tend to give up smoking or at least to moderate consumption.

Among the other discrete explanatory variables, the impact of the occupational variable deserves a specific comment, since the results for the other variables are strictly close to the results previously found in the health economics literature. Whitecollar workers are found to negatively affect the unconditional level of cigarettes with a different impact on probability and on the conditional level. While being a whitecollar reduces the probability to continue smoking, it seems to increase the daily numbers of cigarettes.

Finally, the effect of gender shows that men are more likely to have quit, even though they tend to consume more cigarettes than women. In particular, the estimated parameters highlight that the level of consumption for men is higher by four cigarettes per day than that of women. These results indicate significant differences in gender elasticities and provide support for a closer examination of gender-differentiated smoking patterns.

4.4 Gender differences in smoking behaviour

The role of gender on smoking and the appropriateness of pooling samples are further investigated by estimating all the model presented in Section 3 using separate men and

Table 4 – Likelihood ratio test for gender differences

Model	Log-likelihood value *Pooled*	*Men*	*Women*	*LR*	*df*	*p-value*
BOX-COX D-H DEPENDENT	-118424.47 (42)	-75,128 (40)	-42347.99 (40)	1896.276	38	< 0.001
BOX-COX D-H INDEPENDENT	-118440.56 (42)	-75146.68 (40)	-42347.993 (40)	1891.774	38	< 0.001
D-H DEPENDENT	-120342.25 (42)	-75759.683 (40)	-42912.546 (40)	3340.042	38	< 0.001
D-H INDEPENDENT	-120387.98 (42)	-75799.754 (40)	-42915.562 (40)	3345.328	38	< 0.001
TOBIT	-121826.54 (24)	-76971.537 (23)	-43369.748 (23)	2970.51	22	< 0.001
GENERALIZED TOBIT	-121543.58 (42)	-76949.997 (40)	-42954.151 (40)	3278.864	38	< 0.001
TWO-PART MODEL	-121552.32 (42)	-76954.792 (40)	-42954.29 (40)	3286.476	38	< 0.001

Notes: number of parameters in parentheses

women sub-samples. By means of a LR test, the gender-differentiated estimations of all the models considered are compared to the restricted regression estimated on the pooled sample using the gender dummy variable *Male*. Denoting the maximum log-likelihoods for the men, women and pooled samples as LL_m, LL_w and LL_p, with the corresponding number of parameters k_m, k_w and k_p, the LR statistic is:

$$LR = 2(LL_m + LL_w - LL_p) \qquad (20)$$

which is distributed as a χ^2 with $k_m + k_w - k_p$ degrees of freedom. LR test results are presented in Table 4 and clearly indicates that the hypothesis of equal parameters for men and women is rejected ($p - value < 0.001$) for all the specifications considered. Thus, gender-differentiated estimations are much more informative than those obtained from the pooled sample, suggesting the opportunity of distinguishing the consumption patterns of men and women in modelling cigarette demand.

Table 5 presents results of LR nested specification tests, separately for men and women sub-samples. All the pairwise comparisons for each sample suggest the appropriateness of a bivariate specification for the analysis of smoking behaviour (standard Tobit model is rejected against the standard independent double-hurdle at the 1% significance level) and indicate that the Box-Cox double-hurdle performs better than the other specifications. The only difference between the two samples is that the hypothesis of dependence is not relevant for the female sub-sample, partially contradicting the prediction of the fixed cost model and suggesting that, for women, the benefits of quitting may not be associated with the individual's level of smoking.

Table 5 – Likelihood-ratio tests for nested specifications distinguished by gender

	Men			Women		
Model	*LR*	*df*	*p-value*	*LR*	*df*	*p-value*
Model 1 – Model 2	36.68	1	< 0.001	0.006	1	0.9383
Model 1 – Model 3	1262.68	1	< 0.001	1129.11	1	< 0.001
Model 1 – Model 4	1342.82	2	< 0.001	1135.14	2	< 0.001
Model 4 – Model 5	2343.57	18	< 0.001	908.37	18	< 0.001
Model 1 – Model 6	3643.31	1	< 0.001	1212.32	1	< 0.001
Model 1 – Model 7	3652.90	2	< 0.001	1212.60	2	< 0.001

Table 6 shows the estimated parameters of model (16), separately for men and women, while the calculated elasticities are reported in Table 7. Even though we only comment the elasticities, it is worth noting that many estimated parameters are significant at the 5% significance level and relevant genders-differences in the sign and dimension of the parameters can be pointed out.

As previously, significant habit effects are responsible for the persistence on cigarette consumption and for an increasing the participation probability. The impact of this variable is bigger for men with respect to women, both in the probability to participate and in the conditional level of consumption. It clearly emerges that quitting costs will be more relevant for male smokers, who seems to be less willing to quit. In turn, for the women social interactions assume greater relevance, increasing consumption levels and preventing quitting. The estimated parameters of *Peer_Part* and *Peer_Cons* for women, in fact, reveal a positive and greater impact with respect to the male counterpart.

The discrete variables used in the gender-differentiated estimations highlight some interesting heterogeneous findings. Physical activity has a significant impact on reducing cigarette consumption, mainly for men. Moreover, current health status does not affect cigarette consumption of women, while men in fair/poor health status are characterized by higher consumption levels, coherently with the findings of Yen (2005b).

The dummy for single has a positive effect on smoking participation and consumption decisions. Single are more likely to participate to cigarette consumption, but it is worth noting that single women consume more than men.

Table 6 – Box-Cox double-hurdle estimates distinguished by gender

	Men sub-sample ($n = 31912$)		Women sub-sample ($n = 15865$)	
Variable	*Participation*	*Consumption*	*Participation*	*Consumption*
SMOKING_HABIT	0.1397*** (0.003)	0.1621*** (0.009)	0.0875*** (0.002)	0.0524*** (0.006)
PEER_CONS	–	0.0253*** (0.002)	–	0.0194*** (0.001)
PEER_PART	0.7002*** (0.046)	–	1.1915*** (0.056)	–
HEALTH	–	0.1201*** (0.043)	–	0.0344 (0.035)
CHRONIC_ILLNESS	-0.2885*** (0.024)	-0.1580*** (0.060)	-0.1090*** (0.033)	0.0071 (0.049)
PHYSICAL_ACT	-0.0805*** (0.021)	-0.3186*** (0.042)	-0.0307 (0.027)	-0.1898*** (0.036)
BMI	-0.0333*** (0.003)	0.0539*** (0.006)	-0.0251*** (0.003)	0.0102** (0.005)
AGE	-0.1427*** (0.003)	-0.1352*** (0.009)	-0.0787*** (0.002)	-0.0353*** (0.005)
SINGLE	0.2246*** (0.031)	0.3939*** (0.054)	0.2213*** (0.033)	0.4281*** (0.048)
CHILDREN013	-0.1256*** (0.022)	–	-0.2950*** (0.027)	–
EDUCATION	0.0017 (0.003)	-0.0162*** (0.006)	-0.0170*** (0.004)	-0.0028 (0.005)
PROFESSIONAL	0.0291 (0.049)	0.1022 (0.109)	0.0435 (0.082)	0.1929 (0.123)
WHITECOLLAR	-0.0291 (0.025)	-0.0239 (0.051)	-0.0368 (0.030)	0.0539 (0.043)
UNEMPLOYED	0.1393*** (0.040)	-0.4708*** (0.073)	0.0423 (0.042)	-0.1796*** (0.058)
STUDENT	-0.0186 (0.059)	-1.6952*** (0.103)	0.1575*** (0.067)	-0.7502*** (0.073)
RETIRED	-0.1464*** (0.033)	-1.2280*** (0.090)	0.0052 (0.043)	-0.1800*** (0.064)
NORTH	0.0100 (0.026)	-0.0801 (0.053)	-0.1052*** (0.031)	-0.0528 (0.043)
SOUTH	0.1360*** (0.026)	0.2708 (0.052)	0.0621* (0.034)	-0.1031*** (0.045)
OWNER_OCC	–	-0.2847*** (0.041)	–	-0.1997*** (0.037)
N_EARNERS	–	-0.1784 (0.035)	–	-0.0932*** (0.025)
EC_RESOURCES	–	0.1048*** (0.042)	–	0.0328 (0.037)
QUARTER1	–	-0.0713 (0.051)	–	0.0059 (0.045)
QUARTER2	–	-0.1165*** (0.051)	–	-0.0500 (0.046)
QUARTER3	–	-0.1744*** (0.052)	–	0.0053 (0.045)
CONSTANT	4.2765*** (0.102)	6.7600*** (0.268)	2.8092*** (0.112)	3.9662*** (0.186)
σ	2.4034*** (0.074)		1.6417*** (0.047)	
ρ	-0.2490*** (0.028)		-0.0040 (0.059)	
λ	0.5186*** (0.012)		0.3841*** (0.014)	
Log Likelihood	-75128.342		-42347.99	

Notes: robust standard errors are reported in parentheses.
***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 7 – Elasticities with respect to continuous variables and effects of binary variables

	Men			Women		
Variables	*Probability*	*Cond. level*	*Uncond. level*	*Probability*	*Cond. level*	*Uncond. level*
Continuous variables						
SMOKING_HABIT	2.3675*** (0.0155)	0.6092*** (0.0347)	2.9767*** (0.0134)	0.8897*** (0.0650)	0.1755*** (0.0189)	1.0622*** (0.0548)
PEER_CONS	–	0.0374*** (0.0025)	0.0374 (0.0025)	–	0.0526*** (0.0036)	0.0526*** (0.0036)
PEER_PART	0.2114*** (0.0140)	–	0.2114*** (0.0140)	0.3754*** (0.0177)	–	0.3754*** (0.0177)
BMI	-0.6392*** (0.0155)	0.2084*** (0.0233)	-0.4308*** (0.0134)	-0.2907*** (0.0549)	0.0388** (0.0186)	-0.2519*** (0.0627)
AGE	-5.1664*** (0.3158)	-0.9921*** (0.0649)	-6.1585*** (0.2985)	-1.7114*** (0.1343)	-0.2529*** (0.0360)	-1.9643*** (0.1571)
EDUCATION	0.0129 (0.0517)	-0.0209*** (0.0077)	-0.0080 (0.0374)	-0.2441*** (0.0533)	-0.0044 (0.0084)	-0.2485*** (0.0543)
N_EARNERS	–	-0.0238*** (0.0047)	-0.0238 (0.0047)	–	-0.0153*** (0.0041)	-0.0153*** (0.0041)
Discrete variables						
HEALTH	0.0001	0.4196***	0.2043**	0.0000	0.1606	0.1073
CHRONIC_ILLNESS	-0.1139***	-0.1577**	-1.6297***	-0.0400***	0.0313	-0.3997**
PHYSICAL_ACT	0.0323***	-2.8477***	-0.9455***	-0.0111	-0.8911***	-0.7162**
SINGLE	0.0894***	1.0945***	1.8282***	0.0772***	2.0395***	2.2888***
CHILDREN013	-0.0499***	0.1698*	-0.6054**	-0.1099***	-0.0049	-1.1574**
PROFESSIONAL	0.0116	0.3192	0.3175	0.0157	0.9112**	0.7862**
WHITECOLLAR	-0.0116	-0.0443	-0.1801	-0.0134	0.2517**	0.0252*
UNEMPLOYED	0.0548***	-1.7927***	-0.2012**	0.0152	-0.8301***	-0.4056**
STUDENT	-0.0114	-5.5090***	-2.7629***	0.0549*	-3.3380***	-1.8223***
RETIRED	-0.0594***	-4.0179***	-2.6690	0.0019	-0.8370***	-0.5400**
NORTH	0.0039	-0.2928*	-0.0881	-0.0384***	-0.2481	-0.5667**
SOUTH	0.0482***	0.6764*	1.1217**	0.0225*	-0.4800***	-0.0862*
OWNER_OCC	-0.0002	-1.0000***	-0.4868**	0.0000	-0.9381***	-0.6267**
EC_RESOURCES	0.0001	0.3665***	0.1785**	0.0000	0.1534	0.1025

Notes: Asymptotic standard errors of estimated elasticities are reported in round brackets.
Asterisks indicate levels of significance: *** = 0.01, ** = 0.05 and * = 0.10.

The presence of children aged 0-13 increases the probability to have quit smoking independently from the sex of the smoker, revealing that individuals, as a form of responsibility, tend to cut smoking when small children are present within the family.

Moreover, employment and social status variables have different effects among genders. For men, being a professional or having a whitecollar occupation does not affects neither the participation probability or the consumption level. On the other hand, for the women sub-sample, participation still remains unaffected by the working status variables, while professional and whitecollar women are characterized, conditional on smoking, by higher cigarette consumption levels. Also the effect of education differs between genders. For women, education plays a negative and significant role on the probability, but not on the conditional level of cigarette consumption. The overall effect on the unconditional level is significant and negative, as the effect on probability dominates. This result suggests that while education lowers the probability to continue smoking, it will not reduce consumption levels among smoking women. On the contrary, for men, education is effective in reducing the number of cigarette smoked, while it does not affect the participation probability.

These evidences can be complemented by the analysis of the effects of the geographical characterization of smoking behaviours. These differences are statistically relevant. While the probability to continue smoking for men is higher in the South of Italy, we obtain a positive impact on the level of cigarette consumption for smokers in the South and negative for those of the Northern area. Conversely, the *North* and *South* dummies negatively impact on the women cigarette consumption., indicating that mainly are the women of the centre of Italy to increase smoking consumption.

5 Concluding Remarks

In this paper, the effects of individual and social factors on cigarette demand and on the decision to quit smoking have been analyzed for Italy. The empirical analysis of the impact of addiction on consumption and quitting showed a strong increase of the level of cigarettes and a lower probability of quitting. From the observation that preferences differ between genders, the estimated models for men and women have allowed to draw two relevant conclusions. Firstly, the fixed costs associated with addiction are higher for men that for women. Thus, aiming at reducting cigarette consumption levels, a persistence in reactions can be expected in male smokers. Secondly, social interactions are statistically more relevant in explaining cigarette consumption and the disincentive to quit for women.

Moreover, it emerges that together with addictive individual behaviour and social externalities, demographic variables are responsible for heterogeneous behaviour of Italian smokers. The estimation results, which are strictly close to the findings of recent empirical studies, provides a clear indication of the relevance of individual characteristics in determining both participation and consumption decisions, that have to be considered when designing public health policies.

The results obtained in this paper are in line with the mix of instruments adopted by the 2004 Italian smoking reform (Article 51 of Law 3/2003, effective from January 10 2005), that bans smoking in public places and finances awareness campaigns on the health risks of cigarette consumption. Smoking bans and anti-tobacco advertising may, in fact, be more effective than taxation policies in modifying individual smoking attitudes. In particular, from the results obtained in the present analysis, it emerges that women could be likely more susceptible to give up smoking (or at least to reduce consumption) since they are characterized by lower quitting costs than men and they are found to be more influenced by social interactions. An answer to this intriguing question would, however, require further research on the pattern of smoking behaviour after the reform.

References

Aristei, D. and L. Pieroni (2007) A Double-Hurdle Approach to Modelling Tobacco Consumption in Italy, *Applied Economics*, forthcoming.

Ary, D. and A. Biglan (1988) Longitudinal changes in adolescent cigarette smoking behavior: onset and cessation, *Journal of Behavioral Medicine* **11**, 361-382

Atkinson, A.B., J. Gomulka, and N.H. Stern (1984) Household expenditure on tobacco 1970-1980: evidence from the Family Expenditure Survey, ESRC Programme on Taxation, Incentives, and the Distribution of Income, London School of Economics, Discussion Paper No. 57.

Auld, M.C. (2005) Causal effect of early initiation on adolescent smoking patterns, *Canadian Journal of Economics* **38**, 709-734.

Becker, G.S. (1996) *Accounting for Tastes*, Harvard University Press, Cambridge.

Becker, G.S. and K. Murphy (1988) A Theory of Rational Addiction, *Journal of Political Economy* **96**, 675-701.

Becker, G.S., M. Grossman and K.M. Murphy (1994) An Empirical Analysis of Cigarette Addiction, *American Economic Review* **84**, 396-418.

Blaylock, J.R. and W.N. Blisard (1992) U.S. Cigarette consumption: The Case of Low-Income Women, *American Journal of Agricultural Economics* **74**, 698-705.

Blaylock, J.R. and W.N. Blisard (1993) Wine consumption by US men, *Applied Economics* **24**, 645-651.

Blundell, R. and C. Meghir (1987) Bivariate Alternatives to the Univariate Tobit Model, *Journal of Econometrics* **34**, 179-200.

Brock, W.A., and S.N. Durlauf (2001) Discrete choice with social interactions, *Review of Economic Studies* **68**, 235-260.

Chaloupka, F. (1991) Rational Addictive Behavior and Cigarette Smoking, *Journal of Political Economy* **99**, 722-42.

Cragg, J. (1971) Some statistical models for limited dependent variables with application to the demand for durable goods, *Econometrica* **39**, 829-844.

Garcia, J. and J.M. Labeaga (1996) Alternative Approaches to Modelling Zero Expenditure: An Application to Spanish Demand for Tobacco, *Oxford Bulletin of Economics and Statistics* **58**, 489-506.

Gaviria, A. and S. Raphael (2001) School-based peer effects and juvenile behavior, *Review of Economics and Statistics* **83**, 257-268.

Heckman, J. (1979) Sample selection bias as a specification error, *Econometrica* **47**, 153-161.

Jiménez-Martín, S., J.M. Labeaga and A. López (1998) Participation, heterogeneity and dynamics in tobacco consumption: evidence from cohort data, *Health Economics* 7, 401-414.

Jones, A.M. (1989) A double-hurdle model of cigarette consumption, *Journal of Applied Econometrics* **4**, 23-39.

Jones, A.M. (1992) A note on computation of the double-hurdle model with dependence with an application to tobacco expenditure, *Bulletin of Economic Research,* **44**, 67-74.

Jones, A.M. (1994) Health, addiction, social interaction and the decision to quit smoking, *Journal of Health Economics* **13**, 93-110.

Jones, A.M. (1996) Smoking cessation and health: a comment, *Journal of Health Economics* **15**, 755-759.

Jones, A.M. (1999) Adjustment costs, withdrawal effects, and cigarette addiction, *Journal of Health Economics* **18**, 125-137.

Jones, A.M. and J.M. Labeaga (2003) Individual heterogeneity and censoring in panel data estimates of tobacco expenditure, *Journal of Applied Econometrics* **18**, 157-177.

Jones, A.M. and S.T. Yen (2000) A Box-Cox double-hurdle model, *The Manchester School* **68**, 203-221.

Krauth, B. (2005) Peer effects and selection effects on smoking among Canadian youth, *Canadian Journal of Economics* **38**, 735-757.

Labeaga, J.M. (1993) Individual behaviour and tobacco consumption: a panel data approach, *Health Economics* **2**, 103-112.

Labeaga, J.M. (1999) A double-hurdle rational addiction model with heterogeneity: estimating the demand for tobacco, *Journal of Econometrics* **93**, 49-72.

Manski, C.F. (1993) Identification of endogenous social effects: the reflection problem, *Review of Economic Studies* **60**, 531-542.

Manski, C.F. (1995) *Identification problems in the social sciences,* Harvard University Press, Cambridge.

Maurer, J. and A. Meier (2005) Do the "Joneses" Really Matter? Peer-group vs. Correlated, IFS Working Paper W05/15.

Pohlmeier, W. and V. Ulrich (1995) An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care, *The Journal of Human Resources* **30**, 339-361.

Powell, L.M., J.A. Tauras and H. Ross (2005) The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behaviour, *Journal of Health Economics* **24**, 950-968.

Sacerdote, B. (2001) Peer effects with random assignment: results for Dartmouth roommates, *Quarterly Journal of Economics* **116**, 681-704.

Shmueli, A. (1996) Smoking cessation and health, *Journal of Health Economics* **15**, 751-754.

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

Tobin, J. (1958) Estimation of relationships for limited dependent variables, *Econometrica*, **26**, 24-36.

Yen, S.T. (1993) Working wives and food away from home: the Box-Cox double hurdle model, *American Journal of Agricultural Economics* **75**, 884-895.

Yen, S.T. (1999) Gaussian versus count-data hurdle models: cigarette consumption by women in the US, *Applied Economics Letters* **6**, 73-76.

Yen, S.T. (2005a) Zero observations and gender differences in cigarette consumption, *Applied Economics* **37**, 1839-1849.

Yen, S.T. (2005b) A Multivariate Sample-Selection Model: Estimating Cigarette and Alcohol Demands with Zero Observations, *American Journal of Agricultural Economics* **87**, 453-466.

Yen, S.T. and A.M. Jones (1996) Individual Cigarette Consumption and Addiction: a Flexible Limited Dependent Variable Approach, *Health Economics* **5**, 105-117.